

	Type	L #	Hits	Search Text	DBs	Time Stamp	Comments	Error Definition	Errors
1	BRS	L1	109	integrin adj binding adj (motif or domain)	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/29 08:24			0
2	BRS	L2	3405	rgd	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/29 08:24			0
3	BRS	L3	8888	bone adj (growth or formation)	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/29 08:25			0
4	BRS	L4	32	(1 or 2) same 3	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/29 08:51			0
5	BRS	L5	2	wo-9514714-\$.did.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/29 08:26			0
6	BRS	L6	2	5681353.pn.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/29 08:51			0

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:38:39 ; Search time 21.3158 Seconds
(without alignments)
146.247 Million cell updates/sec

Title: US-09-812-485A-47

Perfect score: 82

Sequence: 1 ERGDNISPFGDQG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	15	10	US-09-812-485A-47
2	82	100.0	15	12	US-10-360-202-10
3	82	100.0	19	12	US-10-360-202-7
4	82	100.0	19	12	US-10-360-202-9
5	82	100.0	20	12	US-10-360-202-8
6	82	100.0	21	12	US-10-360-202-6
7	82	100.0	23	10	US-09-812-485A-49
8	82	100.0	23	12	US-10-360-202-1
9	82	100.0	24	12	US-10-360-202-17
10	82	100.0	24	12	US-10-360-202-18
11	82	100.0	97	10	US-09-812-485A-1
12	82	100.0	509	12	US-09-794-422-34
13	82	100.0	525	10	US-09-814-550-2
14	82	100.0	525	12	US-09-794-422-6
15	82	100.0	525	12	US-10-311-840-1

16	82	100.0	540	12	US-09-794-422-46	Sequence 46, Appl
17	82	100.0	556	12	US-09-794-422-8	Sequence 8, Appl
18	79	96.3	15	12	US-10-360-202-15	Sequence 15, Appl
19	79	96.3	23	12	US-10-360-202-4	Sequence 4, Appl
20	77	93.9	40	10	US-09-812-485A-23	Sequence 23, Appl
21	71	86.6	15	12	US-10-360-202-16	Sequence 16, Appl
22	71	86.6	19	10	US-09-814-550-6	Sequence 6, Appl
23	71	86.6	23	12	US-10-360-202-5	Sequence 5, Appl
24	71	86.6	435	12	US-09-794-422-2	Sequence 2, Appl
25	71	86.6	441	12	US-09-794-422-4	Sequence 4, Appl
26	68	82.9	15	12	US-10-360-202-14	Sequence 14, Appl
27	68	82.9	20	12	US-10-360-202-13	Sequence 13, Appl
28	68	82.9	21	12	US-10-360-202-12	Sequence 12, Appl
29	68	82.9	22	12	US-10-360-202-11	Sequence 11, Appl
30	68	82.9	23	12	US-10-360-202-2	Sequence 2, Appl
31	66	80.5	12	10	US-09-812-485A-42	Sequence 42, Appl
32	61	74.4	23	12	US-10-360-202-3	Sequence 3, Appl
33	60	73.2	15	10	US-09-812-485A-48	Sequence 48, Appl
34	60	73.2	33	10	US-09-812-485A-30	Sequence 30, Appl
35	55	67.1	15	10	US-09-812-485A-46	Sequence 46, Appl
36	49.5	60.4	40	10	US-09-812-485A-24	Sequence 24, Appl
37	49	59.8	30	10	US-09-812-485A-26	Sequence 26, Appl
38	48.5	59.1	35	10	US-09-812-485A-25	Sequence 25, Appl
39	43	52.4	1756	12	US-10-369-433-5843	Sequence 5843, Ap
40	43	52.4	1758	12	US-10-369-433-5842	Sequence 5842, Ap
41	42	51.2	1209	14	US-10-108-605-349	Sequence 349, App
42	41	50.0	1458	14	US-10-054-691-2	Sequence 2, Appl
43	41	50.0	2039	15	US-10-192-584-7	Sequence 7, Appl
44	41	50.0	2042	15	US-10-192-584-6	Sequence 6, Appl
45	40	48.8	447	15	US-10-156-761-10950	Sequence 10950, A

ALIGNMENTS

RESULT 1

US-09-812-485A-47
; Sequence 47, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russel
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
; FILE REFERENCE: BEAR-006CIP
; CURRENT APPLICATION NUMBER: US/09/812,485A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D-00004 peptide
; NAME/KEY: AMIDATION
; LOCATION: 15
US-09-812-485A-47

Query Match 100.0%; Score 82; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGDNISPFGDQG 15

|||||

Db 1 ERGDNISPFGDQG 15

RESULT 2

US-10-360-202-10

; Sequence 10, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-10

Query Match 100.0%; Score 82; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
| | | | | | | | | | | | | | |
DB 1 ERGNDISPFSGDQ 15

RESULT 3

US-10-360-202-7
; Sequence 7, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-7

Query Match 100.0%; Score 82; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
| | | | | | | | | | | | | | |
DB 5 ERGNDISPFSGDQ 19

RESULT 4

US-10-360-202-9
; Sequence 9, Application US/10360202

; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-9

Query Match 100.0%; Score 82; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
| | | | | | | | | | | | | | |
DB 1 ERGNDISPFSGDQ 15

RESULT 5

US-10-360-202-8
; Sequence 8, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-8

Query Match 100.0%; Score 82; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
| | | | | | | | | | | | | | |
DB 4 ERGNDISPFSGDQ 18

RESULT 6

US-10-360-202-6
; Sequence 6, Application US/10360202
; Publication No. US20030186891A1

```

; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-6

Query Match 100.0%; Score 82; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 5 ERGNDISPFSGDQ 19

RESULT 7
US-09-812-485A-49
; Sequence 49, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russell
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
; FILE REFERENCE: BEAR-008CIP
; CURRENT APPLICATION NUMBER: US/09/812,485A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D-00006 peptide
; NAME/KEY: AMIDATION
; LOCATION: 15
US-09-812-485A-49

Query Match 100.0%; Score 82; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 5 ERGNDISPFSGDQ 19

RESULT 8
US-10-360-202-1
; Sequence 1, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-1

Query Match 100.0%; Score 82; DB 12; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 5 ERGNDISPFSGDQ 19

RESULT 9
US-10-360-202-17
; Sequence 17, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-17

Query Match 100.0%; Score 82; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 6 ERGNDISPFSGDQ 20

RESULT 10
US-10-360-202-18
; Sequence 18, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-17
```

; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; FILE REFERENCE: OF VITAMIN D DEFICIENCY
; CURRENT APPLICATION NUMBER: US/10/360,202
; PRIOR FILING DATE: 2003-02-07
; CURRENT FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-18

Query Match 100.0%; Score 82; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 5 ERGNDISPFSGDQ 19

RESULT 11

US-09-812-485A-1
; Sequence 1, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russell
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; FILE REFERENCE: Peptides and Methods of Treating Skeletal Diseases
; CURRENT APPLICATION NUMBER: US/09/812,485A
; CURRENT FILING DATE: 2001-03-19
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptidic compound
US-09-812-485A-1

Query Match 100.0%; Score 82; DB 10; Length 97;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 47 ERGNDISPFSGDQ 61

RESULT 12

US-09-794-422-34
; Sequence 34, Application US/09794422
; Publication No. US20030166239A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Thomas A.
; APPLICANT: De Wet, Jeffrey R.
; APPLICANT: Gowen, Lori C.
; APPLICANT: Hames, Lynn M.

; TITLE OF INVENTION: Mammalian Osteoregulins
; FILE REFERENCE: PC10445
; CURRENT APPLICATION NUMBER: US/09/794,422
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,617
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/234,500
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-422-34

Query Match 100.0%; Score 82; DB 12; Length 509;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 230 ERGNDISPFSGDQ 244

RESULT 13

US-09-814-550-2
; Sequence 2, Application US/09814550
; Patent No. US20020102641A1
; GENERAL INFORMATION:
; APPLICANT: Schiavi, Susan
; APPLICANT: Madden, Stephen
; APPLICANT: Manavalan, Parthasarathy
; APPLICANT: Levine, Michael
; APPLICANT: Jan de Beur, Suzanne
; TITLE OF INVENTION: ONCOGENIC OSTEOMALACIA-RELATED GENE 1
; FILE REFERENCE: 5014US
; CURRENT APPLICATION NUMBER: US/09/814,550
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,786
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/241,598
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-814-550-2

Query Match 100.0%; Score 82; DB 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
Db 246 ERGNDISPFSGDQ 260

RESULT 14

US-09-794-422-6
; Sequence 6, Application US/09794422
; Publication No. US20030166239A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Thomas A.
; APPLICANT: De Wet, Jeffrey R.
; APPLICANT: Gowen, Lori C.
; APPLICANT: Hames, Lynn M.
; TITLE OF INVENTION: Mammalian Osteoregulins
; FILE REFERENCE: PC10445
; CURRENT APPLICATION NUMBER: US/09/794,422
; CURRENT FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: 60/185,617
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/234,500
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-422-6

Query Match 100.0%; Score 82; DB 12; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
|||
Db 246 ERGNDISPFSGDQ 260
|||

RESULT 15

US-10-311-840-1
; Sequence 1, Application US/10311840
; Publication No. US20030175808A1
; GENERAL INFORMATION:
; APPLICANT: KUROKAWA, Tomofumi
; APPLICANT: YAMADA, Takao
; APPLICANT: MORIMOTO, Shigeto
; TITLE OF INVENTION: No. US20030175808A1e1 Protein and its DNA
; FILE REFERENCE: 2738US0P
; CURRENT APPLICATION NUMBER: US/10/311,840
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: PCT/JP01/05263
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: JP 2000-191088
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 1
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Human
US-10-311-840-1

Query Match 100.0%; Score 82; DB 12; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
|||
Db 246 ERGNDISPFSGDQ 260
|||

Search completed: January 29, 2004, 12:42:43
Job time : 21.3158 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:36:28 ; Search time 10.6579 Seconds
(without alignments)
135.349 Million cell updates/sec

Title: US-09-812-485A-47
Perfect score: 82
Sequence: 1 ERGNDISPFSGDQ 15
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	54.9	913	2	AD32885
2	44	53.7	277	2	lysosome (EC 3.2.1
3	43	52.4	444	2	intracellular alka
4	43	52.4	629	2	probable transmem
5	43	52.4	1758	2	protein C48B4.4b
6	43	52.4	1767	2	transport protein
7	42	51.2	420	2	hypothetical prote
8	42	51.2	874	2	leucyl-tRNA synthe
9	42	51.2	1209	2	brhma associated
10	41	50.0	337	2	hypothetical prote
11	41	50.0	886	2	hypothetical prote
12	41	50.0	1166	2	myosin-like protei
13	40	48.8	140	2	En/Spm-like transp
14	40	48.8	237	2	probable exported
15	40	48.8	300	2	hypothetical prote
16	40	48.8	304	2	probable transcrip
17	40	48.8	357	2	probable protein k
18	40	48.8	381	2	hypothetical prote
19	40	48.8	393	2	cellulase egli - s
20	40	48.8	464	2	arylsulfatase (EC
21	40	48.8	577	2	arylsulfatase (EC
22	40	48.8	798	2	hypothetical prote
23	40	48.8	914	2	protein F10G7.2 (i
24	39	47.6	156	2	unknown protein (i
25	39	47.6	269	2	2-keto-4-pentenat
26	39	47.6	269	2	2-keto-4-pentenat
27	39	47.6	271	2	probable hydratase
28	39	47.6	300	2	dihydrodipicolinat
29	39	47.6	347	2	S-adenosylmethioni

30 39 47.6 402 1 F48552 protein kinase (EC
31 39 47.6 416 2 G39200 hypothetical importe
32 39 47.6 507 2 G89908 Catalase [imported
33 39 47.6 535 2 T29430 amidase homolog -
34 39 47.6 564 2 D96017 probable peptidase
35 39 47.6 576 1 A45049 4-alpha-glucanotra
36 39 47.6 612 2 S65196 probable membrane
37 39 47.6 792 2 F71894 iron (III) dicitra
38 39 47.6 1166 2 S37692 probable tumor sup
39 39 47.6 1526 2 JN0598 DNA topoisomerase
40 39 47.6 1528 2 JS0703 DNA topoisomerase
41 39 47.6 2733 2 S15760 genome polyprotein
42 38.5 47.0 469 1 KCHUI interstitial colla
43 38.5 47.0 469 1 KCPGI interstitial colla
44 38.5 47.0 469 1 KCB01 interstitial colla
45 38.5 47.0 827 1 S66094 stage II sporulati

ALIGNMENTS

RESULT 1
D82885
multiple banded antigen homolog UU487 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: D82885
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mini
A:Reference number: A82870
A:Accession: D82885
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-913 <GLA>
A:Cross-references: GB:AB002146; GB:AF222894; NID:G6899479; PIDN:AAF30899.1; GSPDB:GN0012
A:Experimental source: serovar 3; Biovar 1
C:Genetics:
A:Gene: UU487
A:Genetic code: SGC3

Query Match 54.9%; Score 45; DB 2; Length 913;
Best Local Similarity 81.8%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 NDISPFSGDQ 15
||| ||| |
Db 891 NDISPFSGDQ 901

RESULT 2
AD3322
lysosome (EC 3.2.1.17) - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
C:Accession: AD3322
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, P.; Mazur, M.; Goltman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesec
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3322
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-277 <KUP>
A:Cross-references: GB:AB008917; PIDN:AAL51743.1; PID:gl7982481; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME10562
A:Map position: I
C:Keywords: glycosidase; hydrolase

Query Match 53.7%; Score 44; DB 2; Length 277;

Best Local Similarity 66.7%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RGNDISPFGSD 13
||| ||: ||: ||:
Db 252 RGDADINTFAGD 263

RESULT 3
B83891 intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain C
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B83891
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83891
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA805649.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: aprX

Query Match 52.4%; Score 43; DB 2; Length 444;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ERGNDISPFGSDG 14
: ||: ||: ||: ||: ||:
Db 322 DREDDDVAPSSRG 335

RESULT 4
T39285 Probable transmembrane protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000
C;Accession: T39285
R;McDougall, R.C.; Rajadream, M.A.; Barrell, B.G.; Ramsperger, U.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21841
A;Accession: T39285
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-629 <MCD>
A;Cross-references: EMBL:AL096851; PIDN:CAB50971.1; GSPDB:GN000067; SPDB:SPBC1105.08
A;Experimental source: strain 972h-; cosmid c1105
C;Genetics:
A;Gene: SPDB:SPBC1105.08
A;Map position: 2
A;Introns: 67/1
C;Superfamily: Schizosaccharomyces pombe probable transmembrane protein SPBC1105.08

Query Match 52.4%; Score 43; DB 2; Length 629;
Best Local Similarity 72.7%; Pred. No. 37;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 NDISPFGSDGQ 15
| |||| ||||:
Db 48 NTISPFGDGR 58

RESULT 5
F88559 Protein C48B4.4b [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002
C;Accession: F88559
R;anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: F88559
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1758 <STO>
A;Cross-references: GB:chr_III; PIDN:CAA82384.1; PID:g3875025; GSPDB:GN000021; CESP:C48B4
C;Genetics:
A;Gene: C48B4.4b
A;Map position: 3
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 52.4%; Score 43; DB 2; Length 1758;
Best Local Similarity 53.3%; Pred. No. 1.le+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ERGNDISPFGSDGQ 15
| |||| ||||:
Db 42 ENESNDSTPLRGDQ 56

RESULT 6
S60124 transport protein homolog C48B4.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 13-Jan-1996 #sequence_revision 12-Apr-1996 #text_change 02-Feb-2001
C;Accession: S60124; S40724; S40725
R;Kershaw, J.
submitted to the EMBL Data Library, November 1995
A;Reference number: S60124
A;Accession: S60124
A;Molecule type: DNA
A;Residues: 1-1767 <KER>
A;Cross-references: EMBL:Z29117; NID:g439247; PID:g1066912
C;Genetics:
A;Map position: III
A;Introns: 47/1; 112/3; 161/2; 220/2; 543/3; 574/3; 903/2; 1056/1; 1115/3; 1178/3; 1265/3
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C;Keywords: ATP; duplication; nucleotide binding; P-loop; transmembrane protein
F;528-818/Domain: ATP-binding cassette homology <ABCI>
F;645-652/Region: nucleotide-binding motif A (P-loop)
F;764-769/Region: nucleotide-binding motif B
F;1457-1642/Domain: ATP-binding cassette homology <ABC2>
F;1474-1481/Region: nucleotide-binding motif A (P-loop)
F;1586-1591/Region: nucleotide-binding motif B

Query Match 52.4%; Score 43; DB 2; Length 1767;
Best Local Similarity 53.3%; Pred. No. 1.le+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ERGNDISPFGSDGQ 15
| |||| ||||:
Db 42 ENESNDSTPLRGDQ 56

RESULT 7
AE1857 hypothetical protein all0406 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE1857
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1857
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-420 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA872364.1; PID:g17129751; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0406

Query Match 51.2%; Score 42; DB 2; Length 420;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 NDISPFGDG 14
|||:||||
Db 60 NGINPPAGDG 69

RESULT 8
D75307
leucyl-tRNA synthetase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: D75307
R;White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-874 <WHI>
A;Cross-references: GB:AE002050; GB:AE000513; NID:g6459965; PIDN:AAF11720.1; PID:g645997
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2174
A;Map position: 1
C;Superfamily: leucine-tRNA ligase

Query Match 51.2%; Score 42; DB 2; Length 874;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 GNDISPFGDGQ 15
|||:||||
Db 430 GDQSPFSGEGQ 442

RESULT 9
Tl3153
brahma associated protein 155K - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: Tl3153
R;Papoulas, O.; Beek, S.J.; Moseley, S.L.; McCallum, C.M.; Sarte, M.; Shearn, A.; Tamkun
Development 125, 3955-3966, 1998
A;Title: The Drosophila trithorax group proteins BRM, ASH1 and ASH2 are subunits of dist
A;Reference number: Z17610; MUID:98407952; PMID:9735357
A;Accession: Tl3153
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1209 <PAP>
A;Cross-references: EMBL:AF071502; NID:g3378131; PID:g3378132; PIDN:AAC28454.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0002783
A;Map position: 3
C;Function:
A;Description: an activator of Drosophila homeotic genes

Query Match 51.2%; Score 42; DB 2; Length 1209;
Best Local Similarity 63.6%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RGDNDISPFG 12

Db 380 RGDNDMMPIKG 390
||||:|
RESULT 10
S38166
hypothetical protein YKR088c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YKR408
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Apr-2002
C;Accession: S38166; S42017; S39129
R;Baladron, V.; Ballesta, J.P.G.; Bou, G.; del Rey, F.; Estebean, P.F.; Garcia-Cantalejo,
submitted to the Protein Sequence Database, March 1994
A;Reference number: S38158
A;Accession: S38166
A;Molecule type: DNA
A;Residues: 1-337 <BAL>
A;Cross-references: EMBL:Z28313; NID:g486572; PIDN:CAA82167.1; PID:g486573; MIPS:YKR088c
A;Experimental source: strain S288C
R;Garcia-Cantalejo, J.; Baladron, V.; Estebean, P.F.; Santos, M.A.; Bou, G.; Remacha, M.A.
Yeast 10, 231-245, 1994
A;Title: The complete sequence of an 18,002 bp segment of Saccharomyces cerevisiae chromo
A;Reference number: S42009; MUID:94262327; PMID:8203164
A;Accession: S42017
A;Molecule type: DNA
A;Residues: 1-337 <GAR>
A;Cross-references: EMBL:Z27116; NID:g415899; PIDN:CAA81639.1; PID:g415908
A;Experimental source: strain S288C
C;Genetics:
A;Cross-references: SGD:S0001796
A;Map position: 11R
C;Keywords: transmembrane protein
F;96-112/Domain: transmembrane #status predicted <TM1>
F;138-154/Domain: transmembrane #status predicted <TM2>
F;174-190/Domain: transmembrane #status predicted <TM3>
F;290-306/Domain: transmembrane #status predicted <TM4>

Query Match 50.0%; Score 41; DB 2; Length 337;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DNDISPFGDGQ 15
||||:|
Db 28 DNDIMPNSNGQ 39

RESULT 11
G65036
hypothetical protein b2584 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: G65036
R;Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G65036
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-886 <BLAT>
A;Cross-references: GB:AE000344; GB:U00096; NID:g1788927; PIDN:AAC75637.1; PID:g1788938;
A;Experimental source: strain K-12, substrain MG1655

Query Match 50.0%; Score 41; DB 2; Length 886;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 DISPFSGDGQ 15
|||||
Db 693 DISPFEGDNE 702

RESULT 12

S33812
Myosin-like protein ATM - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001
C/Accession: S33812

R/Knight, A.E.; Kendrick-Jones, J.

J. Mol. Biol. 231, 148-154, 1993

A/Title: A myosin-like protein from a higher plant.

A/Reference number: S33812; MUID:93267647; PMID:7684453

A/Accession: S33812

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-1166 <KNI>

A/Cross-references: EMBL:X69505; NID:q297068; PIDN:CAB61875.1; PID:g6491702

C/Superfamily: myosin heavy chain 95f; myosin motor domain homology

C/Keywords: calmodulin binding; coiled coil; nucleotide binding; P-loop

F;168-825/Domain: myosin motor domain homology <MMOT>

F;256-263/Region: nucleotide-binding motif A (P-loop)

Query Match 50.0%; Score 41; DB 2; Length 1166;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 GNDNDISPFSGDG 14

|||||

Db 63 GDSEDSFYSCHG 74

RESULT 13

G84608

En/Spm-like transposon protein [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C/Accession: G84608

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;

eusa, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: G84608

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-140 <STO>

A/Cross-references: GB:AE002093; NID:g4587588; PIDN:AAD25816.1; GSPDB:GN00139

C/Genetics:

A/Gene: At2g22080

A/Map position: 2

Query Match 48.8%; Score 40; DB 2; Length 140;
Best Local Similarity 57.3%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDG 14

|||||

Db 67 EAGNDDEPFGDGG 80

RESULT 14

AF0772

probable exported protein STY2350 [imported] - Salmonella enterica subsp. enterica serov

C/Species: Salmonella enterica subsp. enterica serovar Typhi

A/Note: this species has also been called Salmonella typhi

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C/Accession: AF0772

R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A/Reference number: AB0502; MUID:21534947; PMID:11677608

A/Accession: AF0772

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-237 <PAR>

A/Cross-references: GB:AL513382; PIDN:CAD02500.1; PID:g16503364; GSPDB:GN00176

C/Genetics:

A/Gene: STY2350

Query Match 48.8%; Score 40; DB 2; Length 237;
Best Local Similarity 46.2%; Pred. No. 41;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 RGDNDISPFSGDG 14

|||||

Db 90 KGESFLSPYIGDG 102

RESULT 15

F90274

hypothetical protein porB-1 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C/Accession: F90274

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-y

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.;

arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: F90274

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-300 <KUR>

A/Cross-references: GB:AE006641; NID:gl13814401; PIDN:AAK41453.1; GSPDB:GN00155

C/Genetics:

A/Gene: porB-1

C/Superfamily: pyruvate synthase beta chain

Query Match 48.8%; Score 40; DB 2; Length 300;
Best Local Similarity 53.3%; Pred. No. 53;
Matches 8; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

QY 2 RGDND--ISPFSGDG 14

|||||

Db 90 RGDNDIAVAVWAGDG 104

Search completed: January 29, 2004, 12:40:56

Job time : 11.6579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:35:18 ; Search time 8.28947 Seconds
(without alignments)
85.096 Million cell updates/sec

Title: US-09-812-485A-47
Perfect score: 82
Sequence: 1 ERGNDISPFSGDGQ 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	51.2	823	1 SYL_DEIRA	Q9rsf0 deinococcus
2	41	50.0	337	1 YK68_YEAST	P36164 saccharomyc
3	41	50.0	886	1 YFIQ_ECOLI	P76594 escherichia
4	40	48.8	83	1 RUXX_THEAC	P57670 thermoplasm
5	40	48.8	393	1 GUN1_USTMA	P54424 ustilago ma
6	40	48.8	464	1 ARS_KLEAE	P20713 klebsiella
7	39	47.6	269	1 MHPD_ECOLI	P77608 escherichia
8	39	47.6	300	1 DAPA_MYCTU	Q33295 mycobacteri
9	39	47.6	347	1 QUEA_PSEAE	Q9hxh8 pseudomonas
10	39	47.6	402	1 U120_HSYMV	Q05101 marek's dis
11	39	47.6	576	1 DPEP_SOLTU	Q06801 solanum tub
12	39	47.6	612	1 YP68_YEAST	Q08925 saccharomyc
13	39	47.6	1166	1 L2GI_DROPS	Q08470 drosophila
14	39	47.6	1526	1 TP2A_RAT	P41516 rattus norv
15	39	47.6	1528	1 TP2A_MOUSE	Q01320 mus musculu
16	39	47.6	1953	1 BIGA_SALTY	P21927 salmonella
17	39	47.6	2733	1 RRPB_CVNA5	P16342 murine coro
18	38.5	47.0	469	1 MM01_BOVIN	P28053 bos taurus
19	38.5	47.0	469	1 MM01_HORSE	Q9x825 equus cabal
20	38.5	47.0	469	1 MM01_HUMAN	P03956 homo sapien
21	38.5	47.0	469	1 MM01_PIG	P21692 sus scrofa
22	38.5	47.0	827	1 SP2E_BACSU	P37475 bacillus su
23	38	46.3	263	1 PSBP_CUCSA	Q9slc8 cucumis sat
24	38	46.3	621	1 GIDA_BORBU	P53362 borrelia bu
25	38	46.3	684	1 RPSD_AGR75	P33452 agrobacteri
26	38	46.3	996	1 ATAI_MAKN1	P70083 m sarcoplas
27	37	45.1	224	1 DCL_LYCES	Q42463 lycopersico
28	37	45.1	301	1 PORI_RHOCA	P12143 rhodobacter
29	37	45.1	301	1 TSA2_MOUSE	Q8vic3 mus musculu
30	37	45.1	309	1 TS22_HUMAN	Q8wy24 homo sapien
31	37	45.1	329	1 IPNS_STRJU	P18286 streptomyce
32	37	45.1	352	1 YK50_YEAST	P36151 saccharomyc
33	37	45.1	389	1 UT1_HUMAN	Q13336 homo sapien

RESULT 1
SYL_DEIRA

ID	SYL_DEIRA	STANDARD;	PRT;	823 AA.
AC	Q9RSF0;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Leucyl-tRNA synthetase (EC 6.1.1.4) (Leucine--tRNA ligase) (Leurs)			
GN	LEUS OR DR2174.			
OS	Deinococcus radiodurans.			
OC	Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;			
OC	Deinococcaceae; Deinococcus.			
OX	NCBI_TaxID=1299;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;			
RC	MEDLINE=20036896; PubMed=10567266;			
RA	White O., Eisen J.A., Heidelberg J.P., Hickey E.K., Peterson J.D.,			
RA	Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,			
RA	Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,			
RA	Vamatheva J.J., Lam P., McDonald L., Utterback T., Zaleski C.,			
RA	Makareva K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,			
RA	Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,			
RA	Fraser C.M.;			
RT	"Genome sequence of the radioresistant bacterium Deinococcus			
RT	radiodurans R1.";			
RL	Science 286:1571-1577(1999).			
CC	-I- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +			
CC	diphosphate + L-leucyl-tRNA(Leu).			
CC	-I- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AE002050; AAF11720.1; ALT_INIT.			
DR	TIGR; DR2174;			
DR	HAWAP; MF_00049; -; 1.			
DR	InterPro; IPR002302; Leu-trNASynt1a.			
DR	InterPro; IPR002300; tRNA-synt_1a.			
DR	InterPro; IPR001412; tRNA-synt_1.			
DR	Pfam; PF00113; tRNA-synt_1; 1.			
DR	PRINTS; PR00985; TRNASYNTHLEU.			
DR	TIGRFAMs; TIGR00396; leus bact; 1.			
DR	PROSITE; PS00178; AA_TRNA_Ligase_I; FALSE NEG.			
KW	Aminoacyl-tRNA synthetase; Ligase; ATP-binding;			
KW	Complete proteome.			
FT	SITE 55 65 "HIGH" REGION.			
FT	SITE 590 594 "KMSKS" REGION.			
FT	BINDING 593 593 ATP (BY SIMILARITY).			
SQ	SEQUENCE 823 AA; 92613 MW; 2E6445F39005A110 CRC64;			

ALIGNMENTS

```

Query Match          51.2%; Score 42; DB 1; Length 823;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 GDNDISPFGDQ 15
DB 379 GDQSEPYSGEG 391

RESULT 2
YK68_YEAST
ID YK68_YEAST STANDARD; PRT; 337 AA.
AC P36164;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Hypothetical 38.3 kDa protein in PRF16-SRP40 intergenic region.
GN YKR088C OR YKR408.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=94262327; PubMed=8203164;
RX Garcia-Cantalejo J., Baladron V., Esteban P.F., Santos M.A., Bou G.,
RA Remacha M.A., Revuelta J.L., Ballesta J.P.G., Jimenez A., del Rey F.;
RT "The complete sequence of an 18,002 bp segment of Saccharomycetes
RT cerevisiae chromosome XI contains the HBS1, MRP-L20 and PRP16 genes,
RT and six new open reading frames.";
RL Yeast 10:231-245(1994).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z27116; CAA81639.1; -
DR EMBL; Z28313; CAA82167.1; -
DR PIR; S38166; S38166.
DR SGD; S0001796; YKR088C.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 96 115 POTENTIAL.
FT TRANSMEM 138 162 POTENTIAL.
FT TRANSMEM 173 191 POTENTIAL.
FT TRANSMEM 222 246 POTENTIAL.
FT TRANSMEM 253 271 POTENTIAL.
FT TRANSMEM 287 309 POTENTIAL.
SQ SEQUENCE 337 AA; 38311 MW; 7EA95DD4E5AF77FE CRC64;

Query Match          50.0%; Score 41; DB 1; Length 337;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 DNDISPFGDQ 15
DB 28 DNDIMPNSNGQ 39

RESULT 3
YFIQ_ECOLI
ID YFIQ_ECOLI STANDARD; PRT; 886 AA.
AC P76594; O47320;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yfiQ.
GN YFIQ OR B2584.

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RN SEQUENCE OF 1-612 FROM N.A.
RP STRAIN=K12;
RC Nashimoto H., Saito N.;
RA Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: SOME, IN THE N-TERMINAL, TO SUCCINYL-COA LIGASES.
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0590.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000344; AAC75637.1; -
DR EMBL; D64044; BAA10925.1; -
DR PIR; G65036; G65036.
DR EcoGene; EGI4224; YfiQ.
DR InterPro; IPR003781; CoA_binding.
DR InterPro; IPR000182; GCN5acetyltransf.
DR Pfam; PF00583; Acetyltransf; 1.
DR Pfam; PF02629; CoA_binding; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 506 507 ST -> YA (IN REF. 2).
SQ SEQUENCE 886 AA; 97987 MW; 255944B9E2961251 CRC64;

Query Match          50.0%; Score 41; DB 1; Length 886;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 6 DISPFGDQ 15
DB 693 DISPFEGDNE 702

RESULT 4
RUXX_THEAC
ID RUXX_THEAC STANDARD; PRT; 83 AA.
AC P57670;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative snRNP Sm-like protein.
GN TAI240.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).

```

CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL445067; CAC12364.1; -.
 DR HAMAP; MF 00257; -; 1.
 DR InterPro; IPR006649; snRNP.
 DR InterPro; IPR001163; snRNP_Sm.
 DR Pfam; PF01423; Sm; 1.
 DR ProDom; PD020287; snRNP; 1.
 DR SMART; SM00651; Sm; 1.
 KW Hypothetical protein; Ribonucleoprotein; Complete proteome.
 SQ SEQUENCE 83 AA; 9070 MW; 07D132A75150864D CRC64;

 Query Match 48.8%; Score 40; DB 1; Length 83;
 Best Local Similarity 66.7%; Pred. No. 4.2;
 Matches 10; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

 QY 2 RGDNDI--SPFSGDG 14
 DB 68 RGDNVIFSPSKGDG 82

 RESULT 5
 GUNI_USTMA STANDARD; PRT; 393 AA.
 ID GUNI_USTMA
 AC P54424;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 34, Last annotation update)
 DE Endoglucanase 1 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 1)
 DE (Cellulase 1) (EG 1).
 GN EGL1.
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PBd11.
 RX MEDLINE=96145728; PubMed=8590631;
 RA Schaecker F., Wanner G., Kahmann R.;
 RT "Filament-specific expression of a cellulase gene in the dimorphic
 RT fungus Ustilago maydis."
 RL Biol. Chem. Hoppe-Seyler 376:617-625(1995).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose, lichenin and cereal beta-D-glucans.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: HYPHAL TIP.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN FILAMENTOUS DIKARYON.
 CC -1- PTM: MAY ALSO BE O-GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY K (FAMILY 45 OF GLYCOSYL
 CC HYDROLASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S81598; AAB36147.1; -.
 DR PIR; S59499; S59499.
 DR HSSP; P43316; 2ENG.
 DR InterPro; IPR000334; Glyco_hydro_45.
 DR Pfam; PF02015; Glyco_hydro_45; 1.

DR PROSITE; PS01140; GLYCOSYL_HYDROL_F45; 1.
 KW Cellulose degradation; Hydrolase; Glycosidase; Signal; Glycoprotein.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 393 ENDOGLUCANASE 1.
 FT ACT_SITE 34 34 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 152 152 PROTON DONOR (BY SIMILARITY).
 FT DOMAIN 270 385 ALA/GLY/SER-RICH.
 FT CARBOHYD 343 343 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 393 AA; 39594 MW; 65C753C610CD6AD3 CRC64;

 Query Match 48.8%; Score 40; DB 1; Length 393;
 Best Local Similarity 63.6%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 QY 2 RGDNDISPFG 12
 DB 237 KDDNTISPFG 247

 RESULT 6
 ARS_KLEAE STANDARD; PRT; 464 AA.
 ID ARS_KLEAE
 AC P20713;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
 GN ATSA.
 OS Klebsiella aerogenes.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=28451;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-33.
 RX MEDLINE=90202736; PubMed=2180918;
 RA Murooka Y., Ishibashi K., Yasumoto M., Sasaki M., Sugino H.,
 RA Azakami Y., Yamashita M.;
 RT "A sulfur- and tyramine-regulated Klebsiella aerogenes operon
 RT containing the arylsulfatase (atsA) gene and the atsB gene."
 RL J. Bacteriol. 172:2131-2140(1990).
 CC -1- FUNCTION: AS IS COMMONLY PRODUCED BY SOIL MICROORGANISMS AND PLAYS
 CC AN IMPORTANT ROLE IN THE MINERALIZATION OF SULFATES.
 CC -1- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M31938; AAA25051.1; -.
 DR PIR; B35159; B35159.
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase; 1.
 DR PROSITE; PS00523; SULFATASE_1; 1.
 DR PROSITE; PS00149; SULFATASE_2; 1.
 KW Hydrolase; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 464 ARYLSULFATASE.
 FT ACT_SITE 134 134 POTENTIAL.
 SQ SEQUENCE 464 AA; 51471 MW; C8D09AB9EEF9C627 CRC64;

 Query Match 48.8%; Score 40; DB 1; Length 464;
 Best Local Similarity 53.8%; Pred. No. 29;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

 QY 1 ERGDNDISPFGD 13
 DB 35 DMGYSDISPFGG 47

```

RESULT 7
MHPD_ECOLI
ID MHPD_ECOLI STANDARD; PRT; 269 AA.
AC P77608; P71205; P77045;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 2-keto-4-pentenoate hydratase (EC 4.2.1.-) (2-hydroxypentadienoic acid
DE hydratase).
GN MHPD OR B0350.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Kawamukai M.;
RN [2]
RP Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / CS520;
RA Ferrandez A., Garcia J.L., Diaz E.;
RN Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RN Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 1-6. AND CHARACTERIZATION.
RX MEDLINE=98151237; PubMed=9492273;
RA Pollard J.R., Bugg T.D.;
RT "Purification, characterisation and reaction mechanism of
RT monofunctional 2-hydroxypentadienoic acid hydratase from Escherichia
RT coli.";
RL Eur. J. Biochem. 251:98-106 (1998).
CC -1- COFACTOR: Requires a divalent metal ion for activity, optimum
CC activity is obtained with Mn(2+).
CC -1- ENZYME REGULATION: Inhibited by sodium oxalate.
CC -1- PATHWAY: 3-hydroxyphenylpropionate degradation.
CC -1- SIMILARITY: BELONGS TO THE T0DJ/XVLI/XVLI/HPCG FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D86239; BAA13055.1; -
CC EMBL; Y09555; CAA70750.1; ALT INIT.
CC EMBL; AE000142; AAC73453.1; -
CC EMBL; U73857; AAB18074.1; -
CC EcoGene; EGI4274; mhpD.
CC InterPro; IPR002607; Hydratase decarb.
CC Pfam; PF01689; Hydratase decarb; 1.
CC ProDom; PD003721; Hydratase decarb; 1.
CC Aromatic hydrocarbons catabolism; Lyase; Complete proteome.
KW

FT CONFLICT 202 202 G -> E (IN REF. 1).
SQ SEQUENCE 269 AA; 28890 MW; 34A81A8A4E236358 CRC64;

Query Match 47.6%; Score 39; DB 1; Length 269;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GNDTSPFS 11
|||:||||
DB 89 GDNELPFS 97

RESULT 8
DAPA_MYCTU
ID DAPA_MYCTU STANDARD; PRT; 300 AA.
AC Q33295;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dihydropicolinate synthase (EC 4.2.1.52) (DHDPS).
GN DAPA OR RV2753C OR MT2823 OR MTV002.18C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-aspartate 4-semialdehyde + pyruvate =
CC dihydroadipic acid + 2 H2O.
CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC semialdehyde; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the DHDPS family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL008967; CAA15549.1; -
CC EMBL; AE007110; AAK47142.1; -
CC PIR; H70879; H70879.
CC HSSP; P05640; 1DHP.
CC TIGR; MT2823; -
CC Tuberculist; RV2753C; -

```



```
DR HAMAP; MF 00418; -, 1.
DR InterPro; IPR005263; DapA.
DR InterPro; IPR002220; DHDPs.
DR Pfam; PF00701; DHDPs; 1.
DR PRINTS; PR00146; DHPICSNTHASE.
DR ProDom; PD001859; DHDPs; 1.
DR TIGRFAMs; TIGR00674; dapA; 1.
DR PROSITE; PS00665; DHDPs_1; 1.
DR PROSITE; PS00666; DHDPs_2; 1.
KW Lyase; Diaminopimelate Biosynthesis; Lysine biosynthesis;
KW Complete proteome.
FT ACT_SITE 171 BY SIMILARITY
SQ SEQUENCE 300 AA; 30858 MW; 71CB3228661416367 CRC64;

Query Match 47.6%; Score 39; DB 1; Length 300;
Best Local Similarity 75.0%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 ISFSGDG 14
DB 20 VTFPSGDG 27

RESULT 9
QVEA_PSEAE STANDARD; PRT; 347 AA.
ID QVEA_PSEAE
AC Q9HXH8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S-adenosylmethionine:RNA ribosyltransferase-isomerase (EC 5.-.-.-)
DE (Queuosine biosynthesis protein queA)
GN QVEA OR PA3824.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437317; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: Synthesizes oQ from preQ1 in a single S-
CC adenosylmethionine-requiring step. The ribosyl moiety of AdoMet is
CC transferred and isomerized to the epoxycyclopentane residue of oQ
CC (by similarity).
CC -1- PATHWAY: Queuosine biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE QVEA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004799; AAG07211.1; -.
CC F1R; AB3170; AB3170.
DR HAMAP; MF 00113; -, 1.
DR InterPro; IPR003699; Queuosine_synth.
DR Pfam; PF02547; Queuosine_synth; 1.
DR TIGRFAMs; TIGR00113; queA; 1.
KW Queuosine biosynthesis; Transferase; Isomerase; Complete proteome.

SQ SEQUENCE 347 AA; 38160 MW; 48ECBF074C2BC589 CRC64;

Query Match 47.6%; Score 39; DB 1; Length 347;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 DNDISPSFGD 13
DB 263 DGLKPFSGD 272

RESULT 10
UI20_HSVMG STANDARD; PRT; 402 AA.
ID UI20_HSVMG
AC Q05101;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 44.7 kDa protein kinase (EC 2.7.1.-).
DE US1206 OR US3.
OS Marek's disease herpesvirus (strain GA) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10388;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118245; PubMed=1282282;
RA Sakaguchi M., Urakawa T., Hirayama Y., Miki N., Yamamoto M.,
RA Hirai K.;
RT "Sequence determination and genetic content of an 8.9-kb restriction
RT fragment in the short unique region and the internal inverted repeat
RT of Marek's disease virus type 1 DNA.";
RL Virus Genes 6:365-378(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95313166; PubMed=7831788;
RA Brunovskis P., Velicer L.F.;
RT "The Marek's disease virus (MDV) unique short region:
RT alphaherpesvirus-homologous, fowlpox virus-homologous, and
RT MDV-specific genes.";
RL Virology 206:324-338(1995).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M80595; AAB59895.1; -.
CC ENBL; L22174; AAA64965.1; -.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 102 386 PROTEIN_KINASE.
FT NP_BIND 110 118 ATP (BY SIMILARITY).
FT BINDING 127 127 ATP (BY SIMILARITY).
FT ACT_SITE 218 218 BY SIMILARITY.
SQ SEQUENCE 402 AA; 44715 MW; 93E29494B3572C7B CRC64;

Query Match 47.6%; Score 39; DB 1; Length 402;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ERGDNDISPSGDG 14
| : : : : |
| : : : : |
```

```
Db 69 ESQEDLSPEFNDG 82
RESULT 11
DPEP_SOLUTU STANDARD; PRT; 576 AA.
AC Q06801;
DT 30-MAY-2000 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE 4-alpha-glucanotransferase, chloroplast precursor (EC 2.4.1.25)
DE (Amylomaltase) (Disproportionating enzyme) (D-enzyme).
GN DPEP.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 53-57; 174-183 AND 247-259.
RC STRAIN=cv. May Queen; TISSUE=tuber;
RX MEDLINE=93123262; PubMed=7678257;
RA Takaha T., Yanase M., Okada S., Smith S.M.;
RT "Disproportionating enzyme (4-alpha-glucanotransferase; EC 2.4.1.25)
of potato. Purification, molecular cloning, and potential role in
starch metabolism."
RL J. Biol. Chem. 268:1391-1396(1993).
CC -1- FUNCTION: MAY ACT DURING STARCH BREAKDOWN TO CONVERT SMALL
CC OLIGOSACCHARIDES INTO LARGER MOLECULES UPON WHICH STARCH
CC PHOSPHORYLASE CAN ACT, OR MAY CHANGE THE STRUCTURE OF STARCH
CC MOLECULES AND GRAIN ARCHITECTURE BY MODIFYING CHAIN LENGTH, OR
CC MAY GENERATE FROM STARCH AND GLUCOSE OLIGOSACCHARIDES WHICH
CC CAN SERVE EITHER AS PRIMERS FOR NEW STARCH PHOSPHOENZYME.
CC -1- CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan
CC to a new 4-position in an acceptor, which may be glucose or (1,4)-
CC alpha-D-glucan.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST OR AMYLOPLAST.
CC -1- TISSUE SPECIFICITY: PRESENT IN LEAVES, STEMS, ROOTS, AND STOLONS
CC BUT IS MOST ABUNDANT IN DEVELOPING AND MATURE TUBERS.
CC -1- SIMILARITY: BELONGS TO THE DISPROPORTIONATING ENZYME FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X68664; CAA48630.1; -.
DR PIR; A45049; A45049.
DR HSP; O87172; ICWY.
DR InterPro; IPR003385; Glyco_hydro_77.
DR Pfam; PF02446; 4A_glucoanotrans; 1.
DR TIGRFAMs; TIGR00217; malQ; 1.
KW Transferase; Glycosyltransferase; Carbohydrate metabolism;
KW Amyloplast; Chloroplast; Transit peptide.
FT TRANSIT 1 52 CHLOROPLAST.
FT CHAIN 53 576 4-ALPHA-GLUCANOTRANSFERASE.
SQ SEQUENCE 576 AA; 64950 MW; A0D16F3A546307BB CRC64;
Query Match 47.6%; Score 39; DB 1; Length 576;
Best Local Similarity 58.3%; Pred. No. 54;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 ERGNDISPFSG 12
Db 127 KRNGDGSPYS 138
RESULT 12
YP68_YEAST STANDARD; PRT; 612 AA.
ID YP68_YEAST
AC Q08925;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 68.7 kDa protein in PPQ1-WFALL intergenic region.
GN YPL184C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., Dipaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urrestazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."
RL Nature 387:103-105(1997).
CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC -1- SIMILARITY: TO S.POMBE SPCC737.01C AND NRD1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z73540; CAA37894.1; -.
DR PIR; S65196; S65196.
DR SGD; S0006105; YPL184C.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 4.
DR SMART; SM00360; RRM; 4.
DR PROSITE; PS00102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Hypothetical protein; RNA-binding; Repeat.
FT DOMAIN 201 274 RNA-BINDING (RRM) 1.
FT DOMAIN 292 379 RNA-BINDING (RRM) 2.
FT DOMAIN 431 504 RNA-BINDING (RRM) 3.
FT DOMAIN 522 602 RNA-BINDING (RRM) 4.
FT DOMAIN 6 19 POLY-ASN.
FT DOMAIN 22 28 POLY-SER.
FT DOMAIN 45 48 POLY-SER.
FT DOMAIN 413 419 POLY-ALA.
SQ SEQUENCE 612 AA; 68663 MW; A1B2EB381C08BC8E CRC64;
Query Match 47.6%; Score 39; DB 1; Length 612;
Best Local Similarity 46.2%; Pred. No. 58;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 ERGNDISPFSG 13
Db 144 ERSSNEVSPFDE 156
RESULT 13
L2GL_DROPS
```

ID L2GL_DROPS STANDARD; PRT; 1166 AA.
AC Q08470;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Lethal(2) giant larvae protein (P127).
GN L(2)GL.
OS Drosophila pseudoobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7237;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=93275648; PubMed=8389031;
RA Toerckel I., Hattenstein K., Kalmes A., Schmitt R., Strand D.,
RA Mechler B.M.;
RT "The l(2)gl homologue of Drosophila pseudoobscura suppresses
RT tumorigenicity in transgenic Drosophila melanogaster.";
RL Oncogene 8:1537-1549(1993).
CC -!- FUNCTION: COULD ACT AS TUMOR SUPPRESSOR.
CC -!- SUBUNIT: MAY FORM MULTIMERIC COMPLEXES.
CC -!- SUBCELLULAR LOCATION: CELLULAR MEMBRANE OR INTERCELLULAR
CC MATRIX.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE EPITHELIAL CELLS OF THE
CC DIGESTIVE TRACT AND IN GONADS, IN THE OVARY'S NURSE AND OOCYTE'S
CC FOLLICULAR CELLS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED ABUNDANTLY IN EARLY EMBRYOGENESIS.
CC MODERATE EXPRESSION IS FOUND IN LARVAL AND ADULT STAGES.
CC -!- DISEASE: MUTATIONS IN L(2)GL CAUSE MALIGNANT TRANSFORMATION OF THE
CC OPTIC CENTRES OF THE LARVAL BRAIN AND THE IMAGINAL DISCS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X73259; CNA51715.1; --
DR PIR; S37692;
DR FlyBase; FBgn0012725; Dpse\l(2)gl.
DR InterPro; IPR000664; Lethal2 giant.
DR PRINTS; PR00962; LETHAL2GIANT.
KW Anti-oncogene; Repeat
FT REPEAT 525 539 1.
FT REPEAT 615 629 2.
FT SEQUENCE 1166 AA; 127216 MW; 8155B8AB45BD6AC6 CRC64;
SQ
Query Match 47.6%; Score 39; DB 1; Length 1166;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ERGNDISPFSG 12
DB 286 QRGNDVIVFSG 297
RESULT 14
TP2A_RAT
ID TP2A_RAT STANDARD; PRT; 1526 AA.
AC P41516;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase II, alpha isozyme (EC 5.99.1.3).
GN TOP2A OR TOP2 OR TOP-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=93290677; PubMed=8390253;
RA Park S.H., Yoon J.H., Kwon Y.D., Park S.D.;
RT "Nucleotide sequence analysis of the CDNA for rat DNA topoisomerase
RT II.";
RL Biochem. Biophys. Res. Commun. 193:787-793(1993).
CC -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC MAKES DOUBLE-STRAND BREAKS.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC RELAX ONLY NEGATIVE SUPERCOILS.
CC -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46372; CAA86496.1; --
DR EMBL; Z19552; CAA79611.1; --
DR EMBL; Z29676; --; NOT_ANNOTATED_CDS.
DR HSSP; P06786; IBGW.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR003957; CBFA NFYB topis.
DR InterPro; IPR001241; DNA_topoisom.
DR InterPro; IPR002205; DNA_topoisom.
DR Pfam; PF00204; DNA_gyraseB; 1.
DR Pfam; PF00521; DNA_topoisom; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR PRINTS; PR00615; CCAATSUBUNTA.
DR PRINTS; PR00418; TP12FAMILY.
DR ProDom; PD000742; DNA_topoisom; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00433; TOP2c; 1.
DR SMART; SM00434; TOP4c; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.
FT NP_BIND 159 164 ATP (POTENTIAL).
FT ACT_SITE 803 803 DNA CLEAVAGE (BY SIMILARITY).
FT SEQUENCE 1526 AA; 173220 MW; A1961ABDB1B050F CRC64;
SQ
Query Match 47.6%; Score 39; DB 1; Length 1526;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 3 GNDISPFSGD 13
DB 200 GDMELKPFSGE 210
RESULT 15
TP2A_MOUSE
ID TP2A_MOUSE STANDARD; PRT; 1528 AA.
AC Q01370;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase II, alpha isozyme (EC 5.99.1.3).
GN TOP2A OR TOP2 OR TOP-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=93065194; PubMed=1331984;
RA Adachi N., Miyaike M., Ikeda H., Kikuchi A.;
RT "Characterization of cDNA encoding the mouse DNA topoisomerase II
RL that can complement the budding yeast top2 mutation.";
RL Nucleic Acids Res. 20:5297-5303(1992).
RN [2]
RP SEQUENCE OF 1254-1528 FROM N.A.
RC TISSUE=Lymphoma;
RX MEDLINE=94084643; PubMed=8261398;
RA McPherson J., Brown G.A., Goldenberg G.J.;
RT "Characterization of a DNA topoisomerase IIalpha gene rearrangement
RT in adriamycin-resistant P388 leukemia: expression of a fusion
RT messenger RNA transcript encoding topoisomerase IIalpha and the
RT retinoic acid receptor alpha locus.";
RL Cancer Res. 53:5885-5889(1993).
CC -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC MAKES DOUBLE-STRAND BREAKS.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC RELAX ONLY NEGATIVE SUPERCOILS.
CC -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D12513; BAA02076.1; -
DR EMBL; U01915; AAC52135.1; -
DR PIR; JS0703; JS0703.
DR HSP; P06786; IEGW.
DR MGD; MGI:98790; Top2a.
DR InterPro; IPR003594; ATPbind AtPase.
DR InterPro; IPR003957; CBFA NFYB topis.
DR InterPro; IPR001241; DNA_topoisoiI.
DR InterPro; IPR002205; DNA_topoisoiV.
DR Pfam; PF00284; DNA_gyraseb; 1.
DR Pfam; PF00521; DNA_topoisoiV; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR PRINTS; PR00615; CCAATSUBUNTA.
DR PRINTS; PR00418; TPI2FAMILY.
DR ProDom; PD000742; DNA_topoisoiV; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00433; TOP2c; 1.
DR SMART; SM00434; TOP4c; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.
FT NP_BIND 160 165 ATP (POTENTIAL).
FT ACT_SITE 804 804 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 1528 AA; 172876 MW; 9061778DCA6C756A CRC64;

Query Match 47.6%; Score 39; DB 1; Length 1528;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 GDNDISPFSGD 13
Db 201 GDMELKPFSGE 211

```

Search completed: January 29, 2004, 12:39:07
Job time : 8.28947 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:35:53 ; Search time 22.1053 Seconds
(without alignments)
175.107 Million cell updates/sec

Title: US-09-812-485a-47

Perfect score: 82

Sequence: 1 ERGNDISPFSGDQ 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	82	100.0	525	4 Q9NQ76	Q9NQ76 homo sapien
2	79	96.3	500	6 Q95KG5	Q95KG5 macaca fasc
3	79	96.3	555	6 Q9NQ76	Q9NQ76 macaca fasc
4	79	96.3	555	6 Q9GM13	Q9GM13 macaca fasc
5	71	86.6	433	11 Q924I1	Q924I1 mus musculus
6	71	86.6	435	11 Q9ES02	Q9ES02 rattus norv
7	71	86.6	441	11 Q8K4L6	Q8K4L6 mus musculus
8	71	86.6	443	11 Q8K3V0	Q8K3V0 rattus norv
9	60	73.2	275	4 Q8NC19	Q8NC19 homo sapien
10	45	54.9	207	11 Q8V1K4	Q8V1K4 mus musculus
11	45	54.9	237	11 Q91ZW8	Q91ZW8 mus musculus
12	45	54.9	913	16 Q9PQ03	Q9PQ03 ureaplasma
13	44.5	54.3	793	5 Q9VYV0	Q9VYV0 drosophila
14	44	53.7	277	16 Q8Y1B3	Q8Y1B3 bruceella me
15	44	53.7	277	16 Q8FZN1	Q8FZN1 bruceella su
16	44	53.7	752	4 Q9C091	Q9C091 homo sapien

17	44	53.7	1007	16 Q8EH99	Q8EH99 ehewanella
18	43	52.4	444	16 Q9KBJ7	Q9KBJ7 bacillus ha
19	43	52.4	629	3 Q9Y819	Q9Y819 schizosacch
20	43	52.4	1002	5 Q17315	Q17315 procambarus
21	43	52.4	1020	5 Q17314	Q17314 procambarus
22	42	51.2	82	3 Q01675	Q01675 pneumocysti
23	42	51.2	420	16 Q8YZQ1	Q8YZQ1 anabaena sp
24	42	51.2	430	10 Q9LJ25	Q9LJ25 oryza sativ
25	42	51.2	484	16 Q98HCO	Q98HCO rhizobium l
26	42	51.2	702	11 Q8R384	Q8R384 mus musculu
27	42	51.2	1002	5 Q8MSY3	Q8MSY3 drosophila
28	42	51.2	1228	5 Q9VF03	Q9VF03 drosophila
29	42	51.2	1967	12 Q10378	Q10378 rice yellow
30	41.5	50.6	663	8 Q9B6E5	Q9B6E5 yarrowia li
31	41	50.0	412	17 Q9P9M4	Q9P9M4 pyrococcus
32	41	50.0	423	4 Q8IUP7	Q8IUP7 homo sapien
33	41	50.0	454	16 Q8D6W0	Q8D6W0 vibrio vuln
34	41	50.0	1166	10 Q9SNF0	Q9SNF0 arabidopsis
35	41	50.0	1166	10 Q8RW22	Q8RW22 arabidopsis
36	41	50.0	1166	10 Q9LHE9	Q9LHE9 arabidopsis
37	40	48.8	140	10 Q9SHZ7	Q9SHZ7 arabidopsis
38	40	48.8	237	16 Q8ZSE7	Q8ZSE7 salmoneila
39	40	48.8	300	17 Q97YV0	Q97YV0 sulfobobus
40	40	48.8	304	16 Q91798	Q91798 pseudomonae
41	40	48.8	357	10 Q9S1M8	Q9S1M8 arabidopsis
42	40	48.8	381	16 Q9A7E1	Q9A7E1 caulobacter
43	40	48.8	480	10 Q9XE29	Q9XE29 oryza sativ
44	40	48.8	536	12 Q8VA07	Q8VA07 chikungunya
45	40	48.8	563	10 Q8GT32	Q8GT32 sambucus ni

ALIGNMENTS

RESULT 1

Q9NQ76	PRELIMINARY;	PRT;	525 AA.
ID	Q9NQ76		
AC	Q9NQ76;		
DT	01-OCT-2000 (TReMBLrel. 15, Created)		
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)		
DT	01-OCT-2002 (TReMBLrel. 22, Last annotation update)		
DE	Matrix extracellular phosphoglycoprotein precursor.		
GN	MEPE.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Bone;		
RX	MEDLINE=20399567; PubMed=10945470;		
RA	Rowe P.S.N., De Zoyza P.A., Dong R., Wang H.R., White K.E.,		
RA	Econs M.J., Oudet C.L.;		
RT	"MEPE, a new gene expressed in bone marrow and tumors causing		
RT	osteomalacia.";		
RL	Genomics 67:54-68(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21309068; PubMed=11414762;		
RA	Argiro L., Desbarats M., Glorieux F.H., Ecarot B.;		
RT	"Mepe, the gene encoding a tumor-secreted protein in oncogenic		
RT	hypophosphatemic osteomalacia, is expressed in bone.";		
RL	Genomics 74:342-351(2001).		
DR	EMBL; AJ276396; CAB97250.1; -		
DR	EMBL; AF325916; AAK70343.1; -		
DR	Genew; HGNC:13361; MEPE.		
KW	Signal.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	18	525
SQ	SEQUENCE	525 AA;	58419 MW; 0977CAE871CA9E5 CRC64;
Query Match		100.0%;	Score 82; DB 4; Length 525;
Best Local Similarity		100.0%;	Pred. No. 1.3e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQ 15
 |||||:|||||
 Db 246 ERGNDISPFGDQ 260

RESULT 2

Q95KG5 PRELIMINARY; PRT; 500 AA.

AC Q95KG5; (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 55.6 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Temporal cortex;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 libraries.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB060891; BAB46894.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 500 AA; 55577 MW; 918D265AD8EDC7BC CRC64;

Query Match 96.3%; Score 79; DB 6; Length 500;
 Best Local Similarity 93.3%; Pred. No. 4e-05;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQ 15
 |||||:|||||
 Db 222 ERGNDISPFGDQ 236

RESULT 3

Q9N076 PRELIMINARY; PRT; 555 AA.

AC Q9N076; (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Matrix extracellular phosphoglycoprotein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 libraries.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB046056; BAB01638.1; -;
 SQ SEQUENCE 555 AA; 61979 MW; CCFEIA98ADAL19EE4 CRC64;

Query Match 96.3%; Score 79; DB 6; Length 555;
 Best Local Similarity 93.3%; Pred. No. 4.5e-05;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQ 15
 |||||:|||||
 Db 277 ERGNDISPFGDQ 291

RESULT 4

Q9GM13 PRELIMINARY; PRT; 555 AA.

AC Q9GM13; (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE MEPE protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 libraries.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB050259; BAB17010.1; -;
 SQ SEQUENCE 555 AA; 61950 MW; BD4D47B88038A9E2 CRC64;

Query Match 96.3%; Score 79; DB 6; Length 555;
 Best Local Similarity 93.3%; Pred. No. 4.5e-05;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQ 15
 |||||:|||||
 Db 277 ERGNDISPFGDQ 291

RESULT 5

Q924I1 PRELIMINARY; PRT; 433 AA.

AC Q924I1; (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Matrix extracellular phosphoglycoprotein precursor.
 GN MEPE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21309068; PubMed=11414762;
 RA Argiro L., Desbarats M., Glorieux F.H., Ecarot B.;
 RT "Mepe, the gene encoding a tumor-secreted protein in oncogenic
 RT hypophosphatemic osteomalacia, is expressed in bone.";
 RL Genomics 74:342-351(2001).
 DR EMBL; AF314964; AAK70342.1; -;
 DR MGD; MGI:2137384; Mepe.
 KW Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 433 MATRIX EXTRACELLULAR PHOSPHOGLYCOPROTEIN.
 SQ SEQUENCE 433 AA; 45984 MW; 7CD603CAA8CB41B0 CRC64;

Query Match 86.6%; Score 71; DB 11; Length 433;
 Best Local Similarity 85.7%; Pred. No. 0.00081;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RGDNDISPFGDQ 15
 |||||:|||||
 Db 175 RGDNDISPFGDQ 188

RESULT 6

Q9ES02 PRELIMINARY; PRT; 435 AA.

Q9ES02;
01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Osteoregulin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F344;
RX MEDLINE=20549633; PubMed=10967096;
RA Petersen D.N., Tkalcic G.T., Mansolf A.L., Rivera-Gonzalez R.,
Brown T.A.;
RA "Identification of Osteoblast/Osteocyte Factor 45 (OF45), a Bone-
RT specific cDNA Encoding an RGP-containing Protein That Is Highly
RT Expressed in Osteoblasts and Osteocytes";
RL J. Biol. Chem. 275:36172-36180(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=F344;
RA Hu B., Wang X., Wang Y.;
RT "Rattus norvegicus cDNA sequence highly expressed in A1-5 cell line
RT (identical to osteoregulin).";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF260922; AAC33366.1; -
DR EMBL; AF530558; AAM94403.1; -
SQ SEQUENCE 435 AA; 46515 MW; D587F82968A26BCB CRC64;

Query Match 86.6%; Score 71; DB 11; Length 435;
Best Local Similarity 85.7%; Pred. No. 0.00081;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RGNDISPSPSGDQ 15
Db 169 RGNDVPPFSGDQ 182

RESULT 7
Q8K4L6 PRELIMINARY; PRT; 441 AA.
ID Q8K4L6
AC Q8K4L6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Osteoregulin.
GN MEPE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Gowen L.C., Petersen D.N., Vail A.L., Stock J., Tkalcic G.T.,
RA Simmons H.A., Chidsey-Frink K.L., Ke H., McNeish J., Brown T.A.;
RT "Targeted disruption of the osteoregulin gene results in increased
RT bone density";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF298661; AAM87687.1; -
DR MGD; MGI:2137384; Mepe.
SQ SEQUENCE 441 AA; 46872 MW; AA1947BFE9F2E300 CRC64;

Query Match 86.6%; Score 71; DB 11; Length 441;
Best Local Similarity 85.7%; Pred. No. 0.00083;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RGNDISPSPSGDQ 15
Db 183 RGNDVPPFSGDQ 196

RESULT 8
Q8K3V0 PRELIMINARY; PRT; 443 AA.
ID Q8K3V0
AC Q8K3V0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Osteoregulin-like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F344;
RA Wang X., Hu B., Wang Y.;
RT "Rattus norvegicus cDNA sequence expressed in B4 cell line (possible
RT subtype of osteoregulin).";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF530559; AAM94404.1; -
SQ SEQUENCE 443 AA; 47672 MW; 720EEC4C0718FB14 CRC64;

Query Match 86.6%; Score 71; DB 11; Length 443;
Best Local Similarity 85.7%; Pred. No. 0.00083;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RGNDISPSPSGDQ 15
Db 177 RGNDVPPFSGDQ 190

RESULT 9
Q8NC19 PRELIMINARY; PRT; 275 AA.
ID Q8NC19
AC Q8NC19;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90595.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isoagi T., Ota T., Nishikawa T., Hayaishi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagehara K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075076; BAC11386.1; -
KW Hypothetical protein.
SQ SEQUENCE 275 AA; 31358 MW; AF0B426A671B665C CRC64;

Query Match 73.2%; Score 60; DB 4; Length 275;
Best Local Similarity 91.7%; Pred. No. 0.038;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ERGNDISPSPSG 12
Db 246 ERGNDISPSPSG 257

RESULT 10
Q8VIK4 PRELIMINARY; PRT; 207 AA.
ID Q8VIK4
AC Q8VIK4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SIGNR3.

GN CD209D OR SIGNR3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6;
 RA Park C.G., Steinman R.M.;
 RT "Alternatively Spliced Forms of Mouse DC-SIGN Homologs";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF440280; AAL33584.1; -;
 DR MGD; MGI:2157947; CD209d.
 DR InterPro; IPR002353; AntifreezeII.
 DR IntexPro; IPR001304; Lectin_C.
 DR Pfam; PF00059; lectin_c; 1.
 DR PRINTS; PR00356; ANTI-FREEZEII.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE_LLECTIN_2; 1.
 SQ SEQUENCE 207 AA; 23500 MW; 16FBAEB47ECAD3FB CRC64;

Query Match 54.9%; Score 45; DB 11; Length 207;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GDNDISPFGDGD 14
 |||:|||||
 Db 171 GDEDCAEFGDGD 182

RESULT 11

ID Q91ZW8 PRELIMINARY; PRT; 237 AA.
 AC Q91ZW8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE SIGNR3.
 GN CD209D OR SIGNR3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6;
 RX PubMed=11581173;
 RA Park C.G., Takahara K., Umamoto E., Yashima Y., Matsubara K.,
 RA Matsuda Y., Clausen B.E., Inaba K., Steinman R.M.;
 RT "Five mouse homologues of the human dendritic cell C-type lectin, DC-SIGN";
 RL Int. Immunol. 13:1283-1290(2001).
 DR EMBL; AF373411; AAL13237.1; -;
 DR MGD; MGI:2157947; CD209d.
 DR InterPro; IPR001304; Lectin_C.
 DR Pfam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
 DR PROSITE; PS50041; C-TYPE_LLECTIN_2; 1.
 SQ SEQUENCE 237 AA; 26925 MW; 52E3003C65AF8EF6 CRC64;

Query Match 54.9%; Score 45; DB 11; Length 237;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GDNDISPFGDGD 14
 |||:|||||
 Db 201 GDEDCAEFGDGD 212

RESULT 12

Q9PQ03

ID Q9PQ03 PRELIMINARY; PRT; 913 AA.
 AC Q9PQ03;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE M8A N-terminal paralog.
 GN U0487.
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=134821;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Serovar 3;
 RX MEDLINE=20500219; PubMed=11048724;
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
 RA Cassell G.H.;
 RT "The complete sequence of the mucosal pathogen Ureaplasma urealyticum";
 RL Nature 407:757-762(2000).
 DR EMBL; AB02146; AAF30899.1; -;
 KW Complete proteome.
 SQ SEQUENCE 913 AA; 106010 MW; FE9AD007C44AE0E3 CRC64;

Query Match 54.9%; Score 45; DB 16; Length 913;
 Best Local Similarity 81.8%; Pred. No. 56;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 NDISPFGDGD 15
 |||:|||||
 Db 891 NDISPFGDGD 901

RESULT 13

ID Q9VVY0 PRELIMINARY; PRT; 793 AA.
 AC Q9VVY0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE CG14085 protein.
 GN CG14085.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.A.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,


```

RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003517; AAF49175.1; -
DR FlyBase; FBgn0036859; CG14085.
SQ SEQUENCE 793 AA; 88674 MW; FFC0E3DF3E537C41 CRC64;

Query Match 54.3%; Score 44.5; DB 5; Length 793;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 10; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 3 GDNDISPFSG-DGQ 15
Db 693 GDNDSPFQIGQQ 706

RESULT 14
Q8FZNI PRELIMINARY; PRT; 277 AA.
ID Q8YI83
AC Q8YI83;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Lysozyme M1 precursor (EC 3.2.1.17).
GN BMEI0562.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RP MEDLINE=20020109; PubMed=11756688;
RX Delvecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jadhonki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Hasekorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AB009498; AAL51743.1; -
DR InterPro; IPR02053; Glyco_hydro_25.
DR Pfam; PF01183; Glyco_hydro_25; 1.
DR ProDom; PD004620; Glyco_hydro_25; 1.
DR SMART; SM00641; Glyco_25; 1.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 277 AA; 31766 MW; B81CF4BA63E92BB0 CRC64;

Query Match 53.7%; Score 44; DB 16; Length 277;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RGDNDISPFSGD 13
Db 252 RGDADINTFAGD 263

RESULT 15
Q8FZNI PRELIMINARY; PRT; 277 AA.
ID Q8FZNI

```

```

AC Q8FZNI;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycosyl hydrolase, family 25.
GN BR1447.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RP MEDLINE=22247741; PubMed=12271122;
RX Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.P., Madupu R.,
RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.W., Boyie S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AS014440; AAN30360.1; -
DR TIGR; BR1447; -
KW Hydrolase; Complete proteome.
SQ SEQUENCE 277 AA; 31750 MW; B81CF4BA7E28F7B0 CRC64;

Query Match 53.7%; Score 44; DB 16; Length 277;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RGDNDISPFSGD 13
Db 252 RGDADINTFAGD 263

Search completed: January 29, 2004, 12:40:17
Job time : 24.1053 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:34:18 ; Search time 42.9737 Seconds
(without alignments)
84.952 Million cell updates/sec

Title: US-09-812-485A-49

Perfect score: 126

Sequence: 1 TDLQERGDNDISPFSGQGQFFKD 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	126	100.0	23	AAU93726	Dental product bon
2	126	100.0	23	AAO20379	C-terminal amide
3	126	100.0	97	AAU93681	Dental product bon
4	126	100.0	97	AAO20331	Protein of matrix
5	126	100.0	430	AAV53812	Amino acid sequenc
6	126	100.0	430	AA862669	Truncated phosphat
7	126	100.0	509	AA862922	Human osteoregulin
8	126	100.0	525	AA882920	Human osteoregulin
9	126	100.0	525	AA862689	Phosphatonin polyp

10	126	100.0	525	23	ABB08526	Human phosphatonin
11	126	100.0	525	23	AAE13227	Human oncogenic os
12	126	100.0	540	22	AA882923	Human osteoregulin
13	126	100.0	556	22	AA882921	Human osteoregulin
14	101	80.2	40	23	AAU93703	Dental product bon
15	101	80.2	40	23	AAO20353	Protein of matrix
16	86	68.3	19	23	AAE13230	Human OOMI calcium
17	84	66.7	15	23	AAU93725	Dental product bon
18	84	66.7	15	23	AAO20378	C-terminal amide
19	84	66.7	33	23	AAU93710	Dental product bon
20	84	66.7	33	23	AAO20360	Peptide of matrix
21	82	65.1	15	23	AAU93724	Dental product bon
22	82	65.1	15	23	AAO20377	C-terminal amide
23	81	64.3	435	22	AA882918	Rat osteoregulin
24	81	64.3	441	22	AA882919	Mouse osteoregulin
25	80	63.5	275	22	AAU93507	Human polypeptide
26	75	59.5	15	23	AAU93723	Dental product bon
27	75	59.5	15	23	AAO20376	C-terminal amide
28	73.5	58.3	40	23	AAU93704	Dental product bon
29	73.5	58.3	40	23	AAO20354	Protein of matrix
30	72.5	57.5	30	23	AAU93706	Dental product bon
31	72.5	57.5	30	23	AAO20356	Peptide of matrix
32	72.5	57.5	35	23	AAU93705	Dental product bon
33	72.5	57.5	35	23	AAO20355	Peptide of matrix
34	66	52.4	12	23	AAO20372	Peptide of a calci
35	57	45.2	40	23	AAU93712	Dental product bon
36	57	45.2	40	23	AAO20362	Protein of matrix
37	57	45.2	45	23	AAU93707	Dental product bon
38	57	45.2	45	23	AAO20357	Protein of matrix
39	53	42.1	109	22	AAE10214	Human bone marrow
40	53	42.1	472	23	ABF35556	Human phospholipas
41	53	42.1	1216	23	AAE22860	Human phospholipas
42	53	42.1	1458	23	ABO95555	Human lipase NHL (
43	53	42.1	1458	23	ABO95556	Human lipase NHL (
44	53	42.1	1458	24	AAE34440	Human lipid-associ
45	52	41.3	19	23	AAE11229	Human porin peptid

ALIGNMENTS

RESULT 1
AAU93726
ID AAU93726 standard; peptide; 23 AA.
XX
AC AAU93726;
XX
DT 02-JUL-2002 (first entry)
XX
DE Dental product bone growth enhancing peptide #46.
XX
DE Dental product; toothpaste; mouthwash; dental floss; bone growth;
KW integrin binding motif; RGD; skeletal disease; dental disease; tooth;
KW alveolar bone growth; osteoblast; odontoblast; osteopathic.
XX
OS Synthetic.
XX
XX WO200213775-A1.
XX
PD 21-FEB-2002.
XX
PF 09-AUG-2001; 2001WO-US25101.
XX
PR 16-AUG-2000; 2000US-225879P.
XX
XX (BIGB-) BIG BEAR BIO INC.
XX
PI Yoneda T, Nomizu M, Kumagai Y;
XX
XX WPI; 2002-329525/36.
XX
XX Dental product useful for treating skeletal diseases e.g. dental
PT diseases comprises a base material and a compound comprising specific

PT amino acid in a sequence containing the integrin binding motif -
 XX Claim 7; Page 21; 44pp; English.
 XX
 CC The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.
 CC The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 126; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TDQERGDNDISPSGQGPFD 23
 DB 1 TDQERGDNDISPSGQGPFD 23
 RESULT 2
 ID AAO20379 standard; peptide; 23 AA.
 AC AAO20379;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE C-terminal amidated synthesised peptide D-00006.
 XX
 KW Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 KW glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 KW alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 KW weakness; D00006.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Modified-site 23
 FT /note= "C-terminal amide"
 XX
 PN WO200214360-A1.
 XX
 XX 21-FEB-2002.
 XX
 PF 14-AUG-2001; 2001WO-US25542.
 XX
 PR 16-AUG-2000; 2000US-0641034.
 PR 19-MAR-2001; 2001US-0812485.
 XX
 XX (BIGB-) BIG BEAR BIO INC.
 XX
 XX Kumagai Y, Blacher RW, Yoneda T;
 XX WPI; 2002-291971/33.
 XX
 XX New peptide compound useful for reducing bone loss, is capable of
 PT enhancing bone growth, and comprises an integrin binding motif,
 PT glycosaminoglycan binding motif or a calcium binding motif -
 XX
 XX Example 1; Page 15; 50pp; English.
 XX
 CC The invention relates to a peptide compound capable of enhancing bone
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino
 CC acids are in D- or L- conformation and the sequence comprises a motif
 CC selected from an integrin binding motif, a glycosaminoglycan binding

CC motif and a calcium binding motif. The peptide of the invention is useful
 CC for reducing bone loss and for reducing renal phosphate excretion in an
 CC individual. The peptide is useful for promoting regeneration of alveolar
 CC bone and/or teeth, and increases the number and activity of odontoblasts
 CC /osteoclasts that help form dental tissues. The peptide is also useful
 CC for treating or preventing a condition associated with skeletal loss or
 CC weakness. This sequence represents a C-terminal amidated synthesised
 CC peptide D-00006 of the invention.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 126; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TDQERGDNDISPSGQGPFD 23
 DB 1 TDQERGDNDISPSGQGPFD 23
 RESULT 3
 AAU93681
 ID AAU93681 standard; protein; 97 AA.
 XX
 AC AAU93681;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Dental product bone growth enhancing peptide #1.
 XX
 KW Dental product; toothpaste; mouthwash; dental floss; bone growth;
 KW integrin binding motif; RGD; skeletal disease; dental disease; tooth;
 KW alveolar bone growth; osteoblast; odontoblast; osteopathic.
 XX
 OS Synthetic.
 XX
 PN WO200213775-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 09-AUG-2001; 2001WO-US25101.
 XX
 PR 16-AUG-2000; 2000US-225879P.
 XX
 XX (BIGB-) BIG BEAR BIO INC.
 XX
 PI Yoneda T, Nomizu M, Kumagai Y;
 XX WPI; 2002-329525/36.
 XX
 XX Dental product useful for treating skeletal diseases e.g. dental
 PT diseases comprises a base material and a compound comprising specific
 PT amino acid in a sequence containing the integrin binding motif -
 XX
 XX Disclosure; Page 11; 44pp; English.
 XX
 XX The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.
 CC The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.
 XX
 SQ Sequence 97 AA;
 Query Match 100.0%; Score 126; DB 23; Length 97;
 Best Local Similarity 100.0%; Pred. No. 1.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQERGDNDISPSFGDGPFPKD 23
 |||||
 DB 43 TDQERGDNDISPSFGDGPFPKD 65

RESULT 4
 AAO20331
 ID AAO20331 standard; protein; 97 AA.

AC AAO20331;
 XX

DT 31-MAY-2002 (first entry)

DE Protein of matrix extracellular phosphoglycoprotein containing RGD #1.
 KW Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 KW glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 KW alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 KW weakness; matrix extracellular phosphoglycoprotein.
 XX Unidentified.
 OS
 XX
 PN WO200214360-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 14-AUG-2001; 2001WO-US25542.
 XX
 PR 16-AUG-2000; 2000US-0641034.
 PR 19-MAR-2001; 2001US-0812485.
 XX
 PA (BIGS-) BIG BEAR BIO INC.
 XX
 PI Kumagai Y, Blacher RW, Yoneda T;
 XX
 DR WPI; 2002-291971/33.
 XX
 PT New peptide compound useful for reducing bone loss, is capable of
 PT enhancing bone growth, and comprises an integrin binding motif,
 PT glycosaminoglycan binding motif or a calcium binding motif -
 XX
 PS Disclosure; Page 11; 50pp; English.

XX The invention relates to a peptide compound capable of enhancing bone
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino
 CC acids are in D- or L- conformation and the sequence comprises a motif
 CC selected from an integrin binding motif, a glycosaminoglycan binding
 CC motif and a calcium binding motif. The peptide of the invention is useful
 CC for reducing bone loss and for reducing renal phosphate excretion in an
 CC individual. The peptide is useful for promoting regeneration of alveolar
 CC bone and/or teeth, and increases the number and activity of odontoblasts
 CC /osteoclasts that help form dental tissues. The peptide is also useful
 CC for treating or preventing a condition associated with skeletal loss or
 CC weakness. This sequence represents a protein of a matrix extracellular
 CC phosphoglycoprotein containing an RGD motif of the invention.
 XX
 SQ Sequence 97 AA;

Query Match 100.0%; Score 126; DB 23; Length 97;
 Best Local Similarity 100.0%; Pred. NO. 1.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQERGDNDISPSFGDGPFPKD 23
 |||||
 DB 43 TDQERGDNDISPSFGDGPFPKD 65

RESULT 5
 AAY53812
 ID AAY53812 standard; Protein; 430 AA.
 XX
 AC AAY53812;

XX 22-FEB-2000 (first entry)
 DT Amino acid sequence of a human phosphatonin polypeptide.
 DE
 XX
 KW Human; phosphatonin; Metastatic-tumour Excreted Phosphaturic-Element;
 KW MEPE; Na+-dependent phosphate cotransport; vitamin D metabolism;
 KW bone mineralisation; phosphate metabolism related disease;
 KW hyperphosphatemia; renal osteodystrophy; renal dialysis;
 KW secondary hyperparathyroidism; osteitis fibrosa cystica; hypercalcaemia;
 KW X-linked hypophosphatemic rickets; hereditary hypophosphatemic rickets;
 KW hypomineralised bone lesion; stunted growth; cystic fibrosis;
 KW oncogenic hypophosphatemic osteomalacia; renal phosphate leakage;
 KW renal osteodystrophy; osteoporosis; vitamin D resistant rickets;
 KW end organ resistance; renal Fanconi syndrome; autosomal rickets;
 KW Paget's disease; kidney failure; renal tubular acidosis; sprue.
 XX Homo sapiens.
 OS
 XX
 FH Location/Qualifiers
 FT Modified-site 8..10
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 8..11
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 16..21
 FT /note= "myristoylation site"
 FT Modified-site 40..47
 FT /note= "Tyrosine kinase phosphorylation site"
 FT Modified-site 77..79
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 118..120
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 119..124
 FT /note= "myristoylation site"
 FT Modified-site 139..142
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 143..148
 FT /note= "myristoylation site"
 FT Region 152..154
 FT /note= "cell attachment tripeptide"
 FT Region 161..165
 FT /note= "glycosaminoglycan attachment site"
 FT Modified-site 177..180
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 194..197
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 199..202
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 203..205
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 224..227
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 228..231
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 228..230
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 238..241
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 266..271
 FT /note= "myristoylation site"
 FT Modified-site 291..296
 FT /note= "myristoylation site"
 FT Modified-site 311..313
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 312..314
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 315..320
 FT /note= "myristoylation site"
 FT Modified-site 319..321
 FT /note= "protein kinase C phosphorylation site"
 FT Modified-site 325..328
 FT /note= "Casein kinase II phosphorylation site"
 FT Modified-site 370..373

FT Modified-site /note= "amidation site"
 FT 382..386
 FT /note= "Asu-glycosylation site"
 FT 383..387
 FT /note= "Asu-glycosylation site"
 FT 384..386
 FT /note= "protein kinase C phosphorylation site"
 FT 389..394
 FT /note= "myristoylation site"
 FT 403..405
 FT /note= "protein kinase C phosphorylation site"
 FT 405..408
 FT /note= "cAMP and cGMP dependent protein kinase
 phosphorylation site"
 FT 408..410
 FT /note= "protein kinase C phosphorylation site"
 FT 409..411
 FT /note= "protein kinase C phosphorylation site"
 FT 423..426
 FT /note= "Casein kinase II phosphorylation site"
 FT 425..428
 FT /note= "Casein kinase II phosphorylation site"
 FT 427..430
 FT /note= "Casein kinase II phosphorylation site"
 FT W09960017-A2.
 XX 25-NOV-1999.
 XX 18-MAY-1999; 99WO-EP03403.
 XX 18-MAY-1998; 98GB-0010681.
 XX 04-SEP-1998; 98GB-0019387.
 XX (UNLO) UNIV COLLEGE LONDON.
 XX Rowe P;
 XX WPI: 2000-053262/04.
 XX N-PSDB; AAZ36447.
 XX New polypeptides involved in the regulation of phosphate metabolism
 useful for diagnosing and treating disorders related to phosphate
 metabolism -
 XX Claim 6; Fig 8; 136pp; English.
 XX The present sequence represents a phosphatonin polypeptide (also called
 Metastatic-tumour Excreted Phosphaturic-Element (MEPE)). The level of
 phosphatonin in a subject modulates Na⁺-dependent phosphate cotransport,
 vitamin D metabolism and/or bone mineralization. The phosphatonin
 polypeptides, polynucleotides, vectors and antibodies are used to treat
 phosphate metabolism related disease. They are used for treatment of
 hyperphosphatemia, or renal osteodystrophy, hyperphosphatemia in renal
 dialysis/pre-dialysis, secondary hyperparathyroidism or osteitis fibrosa
 cystica, or X-linked hypophosphatemic rickets, hereditary
 hypophosphatemic rickets with hypocalcemia (HHRH), hypomineralised
 bone lesions, stunted growth in juveniles, oncogenic hypophosphatemic
 osteomalacia, renal phosphate leakage, renal osteodystrophy,
 osteoporosis, vitamin D resistant rickets, end organ resistance, renal
 Fanconi syndrome, autosomal rickets, Paget's disease, kidney failure,
 renal tubular acidosis, cystic fibrosis or sprue. The polypeptide
 may also be used to manufacture combined preparations for simultaneous
 separate or sequential use for the treatment of phosphate metabolism
 disorders. A transformed osteoblast or bone cell line capable of
 phosphatonin overexpression is useful for the production of
 phosphatonin.
 XX Sequence 430 AA;
 Query Match 100.0%; Score 126; DB 21; Length 430;
 Best Local Similarity 100.0%; Pred. No. 5.8e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQERGDNDISPFSGDGQPFKD 23
 |||||
 Db 147 TDQERGDNDISPFSGDGQPFKD 169
 RESULT 6
 AAB62669
 ID AAB62669 standard; Protein; 430 AA.
 XX AAB62669;
 AC AAB62669;
 XX 23-JUL-2001 (first entry)
 DT Truncated phosphatonin polypeptide (truncated MEPE).
 XX Metastatic-tumour excreted phosphaturic element; MEPE; phosphatonin;
 KW phosphate; vitamin-D; skeletal formation; mineralization; truncated;
 KW osteopathic; antigout; cytostatic; human.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 161..192
 FT /note= "the amino acid residues in this region are
 indicated incorrectly in the sequence provided
 in the sequence listing"
 FT
 FT
 FT
 XX W0200132878-A2.
 XX 10-MAY-2001.
 XX 31-OCT-2000; 2000WO-BP10747.
 XX 04-NOV-1999; 99US-0434185.
 PR 08-NOV-1999; 99GB-0026424.
 XX (UNLO) UNIV COLLEGE LONDON.
 XX Rowe P;
 PI
 XX WPI: 2001-343487/36.
 DR N-PSDB; AAF83759.
 XX New phosphatonin polypeptide a regulator of phosphate metabolism, for
 diagnosing and treating disorders of phosphate, vitamin-D metabolism,
 skeletal formation e.g. osteoporosis, Paget's disease, gout -
 XX Claim 4; Fig 8; 135pp; English.
 XX The invention relates to a novel human protein, metastatic-tumour
 excreted phosphaturic element (MEPE) or phosphatonin (modulator of
 phosphate and vitamin-D metabolism). The phosphatonin polypeptides,
 polynucleotides and specific antibodies are useful for treating a
 disorder of phosphate or vitamin D metabolism, skeletal formation and
 mineralization. Phosphatonins are used to treat hyperphosphatemia, renal
 osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica
 or gout. It is used to prepare a medicament for treating X-linked
 hypophosphatemic rickets, hereditary hypophosphatemic rickets with
 hypocalcemia (HHRH), hypomineralized bone lesions, stunted growth in
 juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate
 leakage, renal osteodystrophy, osteoporosis, vitamin-D resistant rickets,
 end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's
 disease, kidney failure, renal tubular acidosis, cystic fibrosis or
 sprue. Phosphatonin polynucleotides are useful as molecular weight
 markers on Southern gels, as diagnostic probes for detecting the presence
 of a specific mRNA. Phosphatonin polypeptides are also useful for
 identifying agonists and antagonists, compounds which bind to
 phosphatonin and drug candidates for therapy of phosphate metabolism
 disorders. The present sequence represents a truncated form of
 phosphatonin (MEPE).
 XX Sequence 430 AA;

Query Match 100.0%; Score 126; DB 22; Length 430;
 Best Local Similarity 100.0%; Pred. No. 5.8e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQERGDNDISPFSGDGPFPKD 23
 |||||
 DB 147 TDQERGDNDISPFSGDGPFPKD 169

RESULT 7
 AAB82922
 ID AAB82922 standard; Protein; 509 AA.

XX AC AAB82922;

DT 21-DEC-2001 (first entry)

DE Human osteoregulin (mature polypeptide).

XX Osteoregulin; human; bone; homeostasis; adipose; calcification;
 KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
 KW therapy.

XX Homo sapiens.

XX EP1130098-A2.

XX 05-SEP-2001.

XX 27-FEB-2001; 2001EP-0301768.

XX 29-FEB-2000; 2000US-185617P.

PR 22-SEP-2000; 2000US-234500P.

XX (PFIZ) PFIZER PROD INC.

XX Brown TA, De Wet JR, Gowen LC, Hames LM;

XX WPI; 2001-604111/69.

DR N-PSDB; AAH26810.

XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,
 PT adiposity and calcification of atherosclerotic plaques comprises
 PT measuring the activity of osteoregulin -

PS Claim 1; Page 54-55; 90pp; English.

XX The present sequence is that of human osteoregulin mature
 CC polypeptide, i.e. lacking an N-terminal signal sequence.

CC Osteoregulin is a novel protein which plays a role in regulating
 CC bone homeostasis, adiposity, and the calcification of

CC atherosclerotic plaques. 2 splice variants of human osteoregulin
 CC were identified (see also AAB82923). The invention provides novel

CC osteoregulin proteins, nucleic acids which encode them, vectors,
 CC antibodies, host cells which express heterologous osteoregulins, and
 CC animal cells and mammals with a targeted disruption of an

CC osteoregulin gene. The invention also provides screening assays
 CC to identify modulators of osteoregulin activity as well as methods

CC of treating mammals for diseases or disorders associated with
 CC osteoregulin activity. The modulators of activity may be useful

CC in the manufacture of a medicament for, as well as for treating, a
 CC mammal in need of regulation of bone mass and/or density, adiposity,

CC vascular flexibility, and/or atherosclerotic plaque calcification
 CC (claimed), for treating and preventing osteoporosis, and for

CC stimulating bone repair and regeneration.

XX Sequence 509 AA;

Query Match 100.0%; Score 126; DB 22; Length 509;
 Best Local Similarity 100.0%; Pred. No. 7.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQERGDNDISPFSGDGPFPKD 23
 |||||
 DB 226 TDQERGDNDISPFSGDGPFPKD 248

RESULT 8
 AAB82920

ID AAB82920 standard; Protein; 525 AA.

XX AC AAB82920;

DT 21-DEC-2001 (first entry)

DE Human osteoregulin.

XX Osteoregulin; human; bone; homeostasis; adipose; calcification;
 KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
 KW therapy.

XX Homo sapiens.

XX Key

PH Location/Qualifiers

FT Peptide 1..37

FT /label= Signal_peptide

FT Protein 38..525

FT /label= Mature_protein

XX EP1130098-A2.

XX 05-SEP-2001.

XX 27-FEB-2001; 2001EP-0301768.

XX 29-FEB-2000; 2000US-185617P.

PR 22-SEP-2000; 2000US-234500P.

XX (PFIZ) PFIZER PROD INC.

XX Brown TA, De Wet JR, Gowen LC, Hames LM;

XX WPI; 2001-604111/69.

DR N-PSDB; AAH26808.

XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,
 PT adiposity and calcification of atherosclerotic plaques comprises
 PT measuring the activity of osteoregulin -

PS Claim 1; Page 45-47; 90pp; English.

XX The present sequence is that of human osteoregulin, a novel protein
 CC which plays a role in regulating bone homeostasis, adiposity, and
 CC the calcification of atherosclerotic plaques. The sequence is
 CC predicted from the nucleotide sequence (see AAH26808) of isolated

CC osteoblast cDNA. A splice variant of human osteoregulin was also
 CC identified (see AAB82921). The invention provides novel

CC osteoregulin proteins, nucleic acids which encode them, vectors,
 CC antibodies, host cells which express heterologous osteoregulins, and
 CC animal cells and mammals with a targeted disruption of an

CC osteoregulin gene. The invention also provides screening assays
 CC to identify modulators of osteoregulin activity as well as methods

CC of treating mammals for diseases or disorders associated with
 CC osteoregulin activity. The modulators of activity may be useful

CC in the manufacture of a medicament for, as well as for treating, a
 CC mammal in need of regulation of bone mass and/or density, adiposity,

CC vascular flexibility, and/or atherosclerotic plaque calcification
 CC (claimed), for treating and preventing osteoporosis, and for

CC stimulating bone repair and regeneration.

XX Sequence 525 AA;

Query Match 100.0%; Score 126; DB 22; Length 525;
 Best Local Similarity 100.0%; Pred. No. 7.3e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQLQERGDNDISPFSGDGPFFKD 23
 |||||
 Db 242 TDQLQERGDNDISPFSGDGPFFKD 264

RESULT 9
 AAB62689
 ID AAB62689 standard; Protein; 525 AA.

XX AC AAB62689;

XX DT 23-JUL-2001 (first entry)

XX DE Phosphatonin polypeptide (MEPE).

XX KW Metastatic-tumour excreted phosphaturic element; MEPE; phosphatonin;
 KW phosphate; vitamin-D; skeletal formation; mineralization; truncated;
 KW osteopathic; antigout; cytostatic; human.

XX OS Homo sapiens.

XX PN WO200132878-A2.

XX PD 10-MAY-2001.

XX PF 31-OCT-2000; 2000WO-EP10747.

XX PR 04-NOV-1999; 99US-0434185.

XX PR 08-NOV-1999; 99GB-0026424.

XX PA (UNLO) UNIV COLLEGE LONDON.

XX PI Rowe P;

XX DR WPI; 2001-343487/36.

XX DR N-PSDB; AAF83764.

XX PT New phosphatonin polypeptide a regulator of phosphate metabolism, for
 PT diagnosing and treating disorders of phosphate, vitamin-D metabolism,
 PT skeletal formation e.g. osteoporosis, Paget's disease, gout

XX PS Claim 4; Page 133-134; 135pp; English.

XX CC The invention relates to a novel human protein, metastatic-tumour
 CC excreted phosphaturic element (MEPE) or phosphatonin (modulator of
 CC phosphate and vitamin-D metabolism). The phosphatonin polypeptides,
 CC polynucleotides and specific antibodies are useful for treating a
 CC disorder of phosphate or vitamin D metabolism, skeletal formation and
 CC mineralization. Phosphatons are used to treat hyperphosphatemia, renal
 CC osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica
 CC or gout. It is used to prepare a medicament for treating X-linked
 CC hypophosphatemic rickets, hereditary hypophosphatemic rickets with
 CC hypercalcaemia (HHRH), hypomineralized bone lesions, stunted growth in
 CC juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate
 CC leakage, renal osteodystrophy, osteoporosis, vitamin-D resistant rickets,
 CC end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's
 CC disease, kidney failure, renal tubular acidosis, cystic fibrosis or
 CC sprue. Phosphatonin polynucleotides are useful as molecular weight
 CC markers on Southern gels, as diagnostic probes for detecting the presence
 CC of a specific mRNA. Phosphatonin polypeptides are also useful for
 CC identifying agonists and antagonists, compounds which bind to
 CC phosphatonin and drug candidates for therapy of phosphate metabolism
 CC disorders. The present sequence represents the amino acid sequence of
 CC the entire phosphatonin (MEPE).

XX SQ Sequence 525 AA;

Query Match 100.0%; Score 126; DB 22; Length 525;

Best Local Similarity 100.0%; Pred. No. 7.3e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQLQERGDNDISPFSGDGPFFKD 23

Db 242 TDQLQERGDNDISPFSGDGPFFKD 264
 |||||

RESULT 10

ABB08526
 ID ABB08526 standard; protein; 525 AA.

XX AC ABB08526;

XX DT 23-MAY-2002 (first entry)

XX DE Human phosphatonin.

XX KW Human phosphatonin; cytostatic; antidiabetic; antiinflammatory;
 KW hyperphosphatemia; arteriosclerosis; heart failure;
 KW diabetic renal disease; kidney failure; cystic fibrosis.

XX OS Homo sapiens.

XX PN WO200198495-A1.

XX PD 27-DEC-2001.

XX PF 20-JUN-2001; 2001WO-JP05263.

XX PR 21-JUN-2000; 2000JP-0191088.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Kurokawa T, Yamada T, Morimoto S;

XX WPI; 2002-139791/18.

XX DR N-PSDB; ABA99159, ABA99160.

XX PT Phosphatonin of human origin and DNA encoding it for diagnosis and
 PT treatment of diseases associated with disorders of phosphate
 PT metabolism, e.g., hyperphosphatemia, arteriosclerosis, heart failure,
 PT diabetic renal disease and kidney failure

XX PS Claim 1; Fig 1-4; 130pp; Japanese.

XX CC This invention relates to human phosphatonin which has
 CC the activity of lowering blood phosphate and increasing urinary
 CC phosphate. The proteins are cytostatic, antidiabetic and
 CC antiinflammatory in their action. Phosphatonin down-regulates
 CC sodium-dependent phosphate transport in the kidney, it down-regulates
 CC 25-hydroxy-vitamin D3-1alpha-hydroxylase in the kidney and up-regulates
 CC 25-hydroxy-vitamin D3-24-hydroxylase in the kidney. Phosphatonin is
 CC useful in the diagnosis, treatment and prevention of phosphate
 CC metabolism related diseases such as hyperphosphatemia, arteriosclerosis,
 CC heart failure, diabetic renal disease, kidney failure, acute coronary
 CC disease and cystic fibrosis. This sequence represents human
 CC phosphatonin.

XX SQ Sequence 525 AA;

Query Match 100.0%; Score 126; DB 23; Length 525;

Best Local Similarity 100.0%; Pred. No. 7.3e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDQLQERGDNDISPFSGDGPFFKD 23

|||||

Db 242 TDQLQERGDNDISPFSGDGPFFKD 264

RESULT 11

AAE13227
 ID AAE13227 standard; Protein; 525 AA.

XX AC AAE13227;

XX DT 12-FEB-2002 (first entry)

DE Human oncogenic osteomalacia-related protein-1 (OOM-1).

XX Human; oncogenic osteomalacia-related protein 1; OOM1; gene therapy;

XX phosphate homeostasis-related disease; rickets; osteomalacia; cardiac;

KW rhabdomyolysis; cardiomyopathy; tumoural calcinosis; renal failure;

KW bone mineralisation; phosphaturia; cellular process.

XX Homo sapiens.

OS

XX

XX Key Location/Qualifiers

FT Peptide 1..16

FT /label= Signal_peptide

FT Protein 17..525

FT /note= "Mature oncogenic osteomalacia-related protein-1 (OOM-1)"

FT Domain 130..142

FT /note= "Calcium binding motif"

FT Domain 235..258

FT /note= "Calcium binding motif"

FT Region 247..249

FT /note= "Cell attachment site"

FT Domain 264..275

FT /note= "Calcium binding motif"

FT Domain 412..424

FT /note= "Calcium binding motif"

FT Modified-site 442..455

FT /note= "Potential PKA phosphorylation site"

FT Modified-site 477..481

FT /note= "Glycosylation site"

FT Domain 500..503

FT /note= "Bacterial extracellular solute-binding protein motif"

FT

XX WO200172826-A2.

PN

XX

PD 04-OCT-2001.

XX

XX 22-MAR-2001; 2001WO-US09289.

XX

XX 24-MAR-2000; 2000US-191786P.

PR

PR 19-OCT-2000; 2000US-241598P.

XX

XX (GENZ) GENZYME CORP.

PA (UYJO) UNIV JOHNS HOPKINS.

XX

XX Schiavi S, Madden S, Manavalan P, Levine MDM, Jan De Beur S;

PI

XX WPI; 2002-010720/01.

DR

XX

XX New polynucleotide encoding a polypeptide useful for identifying

PT polynucleotide expression or polypeptide activity modulators used for

PT treating oncogenic osteomalacia, comprises the oncogenic

PT osteomalacia-related gene -

XX

XX Claim 2; Fig 2; 65pp; English.

PS

XX The invention relates to oncogenic osteomalacia-related protein 1

CC designated OOM1 (oncogenic osteomalacia-related factor) and its DNA

CC molecule. OOM1 protein is useful for treating a phosphate homeostasis-

CC related disease such as X-linked hypophosphataemia rickets, oncogenic

CC osteomalacia, rhabdomyolysis, cardiomyopathy, tumoural calcinosis and

CC renal failure. OOM1 proteins and DNA's are useful for modulating the

CC phenotype of a neoplastic cell associated with oncogenic osteomalacia;

CC modulating bone mineralisation; treating a disease characterised by

CC abnormal bone mineralisation; and treating an oncogenic osteomalacia-

CC associated symptom such as hypophosphataemia, phosphaturia, low serum

CC concentrations of 1,25-dihydroxy vitamin D or osteomalacia. OOM1 is

CC useful for modulating renal phosphate transport, which involves altering

CC the activity of OOM1 within a cell and then administering serine protease

CC 11 to the cell. OOM1 DNA is useful for generating non-human transgenic

CC animal models, and for searching and identifying single nucleotide

CC polymorphisms which are mutants, variants of the gene in human

CC population. OOM1 is useful as an immunogen to produce antibodies against

CC OOM1 and in vitro assays to screen for agents or drugs, which modulate

CC cellular processes. The present sequence is human oncogenic

CC osteomalacia-related protein-1 (OOM-1).

CC Note: The present sequence states that this sequence is encoded by the

CC DNA sequence shown in SEQ ID NO: 1 (AAD21889) of the specification.

CC However this does not appear to be the case.

XX

XX Query Match 100.0%; Score 126; DB 23; Length 525;

XX Best Local Similarity 100.0%; Pred. No. 7.3e-11;

XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 1 TDLQERGDNDISPFSGDGPFPKD 23

DB 242 TDLQERGDNDISPFSGDGPFPKD 264

XX

XX RESULT 12

XX AAB82923

XX ID AAB82923 standard; Protein; 540 AA.

XX

XX AC AAB82923;

XX

XX DT 21-DEC-2001 (first entry)

XX

XX DE Human osteoregulin (mature polypeptide).

XX

XX KW Osteoregulin; human; bone; homeostasis; adipose; calcification;

XX KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;

XX KW therapy.

XX

XX OS Homo sapiens.

XX

XX PN EP1130098-A2.

XX

XX PD 05-SEP-2001.

XX

XX PF 27-FEB-2001; 2001EP-0301768.

XX

XX PR 29-FEB-2000; 2000US-185617P.

PR

PR 22-SEP-2000; 2000US-234500P.

XX

XX (PFIZ) PFIZER PROD INC.

XX

XX Brown TA, De Wet JR, Gowen LC, Hames LM;

PI

XX WPI; 2001-604111/69.

DR

DR N-PSDB; AAH26811.

XX

XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,

PT adiposity and calcification of atherosclerotic plaques comprises

PT measuring the activity of osteoregulin -

XX

XX Claim 1; Page 59-61; 90pp; English.

PS

XX The present sequence is that of human osteoregulin mature

CC polypeptide, i.e. lacking an N-terminal signal sequence.

CC Osteoregulin is a novel protein which plays a role in regulating

CC bone homeostasis, adiposity, and the calcification of

CC atherosclerotic plaques. 2 Splice variants of human osteoregulin

CC were identified (see also AAB82922). The invention provides novel

CC osteoregulin proteins, nucleic acids which encode them, vectors,

CC antibodies, host cells which express heterologous osteoregulins, and

CC animal cells and mammals with a targeted disruption of an

CC osteoregulin gene. The invention also provides screening assays

CC to identify modulators of osteoregulin activity as well as methods

CC of treating mammals for diseases or disorders associated with

CC osteoregulin activity. The modulators of activity may be useful

CC in the manufacture of a medicament for, as well as for treating, a

CC mammal in need of regulation of bone mass and/or density, adiposity,

CC vascular flexibility, and/or atherosclerotic plaque calcification

CC (claimed), for treating and preventing osteoporosis, and for
 CC stimulating bone repair and regeneration.

SQ Sequence 540 AA;
 Query Match 100.0%; Score 126; DB 22; Length 540;
 Best Local Similarity 100.0%; Pred. No. 7.6e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDLQERGDNDISPSFGDQPFKD 23
 |||||
 Db 257 TDLQERGDNDISPSFGDQPFKD 279

RESULT 13
 AAB82921
 ID AAB82921 standard; Protein; 556 AA.

XX AC AAB82921;

XX DT 21-DEC-2001 (first entry)

XX DE Human osteoregulin.

XX KW Osteoregulin; human; bone; homeostasis; adipose; calcification;
 XX KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
 XX KW therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Peptide 1..37

XX FT /label= Signal_peptide

XX FT Protein 38..556

XX FT /label= Mature_protein

XX FT Misc-difference 57

XX FT /note= "encoded by TCA"

XX PN EPI130098-A2.

XX PD 05-SEP-2001.

XX PF 27-FEB-2001; 2001EP-0301768.

XX PR 29-FEB-2000; 2000US-185617P.

XX PR 22-SEP-2000; 2000US-234500P.

XX PA (PFIZ) PFIZER PROD INC.

XX PI Brown TA, De Wet JR, Gowen LC, Hames LM;

XX DR WPI; 2001-604111/69.

XX DR N-PSDB; AAB26809.

XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,
 XX adiposity and calcification of atherosclerotic plaques comprises
 XX measuring the activity of osteoregulin -

XX PS Claim 1; Page 48-49; 90pp; English.

XX The present sequence is that of human osteoregulin, a novel protein
 XX which plays a role in regulating bone homeostasis, adiposity, and
 XX the calcification of atherosclerotic plaques. The sequence is
 XX predicted from the nucleotide sequence (see AAB26809) of isolated
 XX osteoblast cDNA. A splice variant of human osteoregulin was also
 XX identified (see AAB82920). The invention provides novel
 XX osteoregulin proteins, nucleic acids which encode them, vectors,
 XX antibodies, host cells which express heterologous osteoregulins, and
 XX animal cells and mammals with a targeted disruption of an
 XX osteoregulin gene. The invention also provides screening assays
 XX to identify modulators of osteoregulin activity as well as methods
 XX of treating mammals for diseases or disorders associated with
 XX osteoregulin activity. The modulators of activity may be useful

CC in the manufacture of a medicament for, as well as for treating, a
 CC mammal in need of regulation of bone mass and/or density, adiposity,
 CC vascular flexibility, and/or atherosclerotic plaque calcification
 CC (claimed), for treating and preventing osteoporosis, and for
 CC stimulating bone repair and regeneration.

XX SQ Sequence 556 AA;

Query Match 100.0%; Score 136; DB 22; Length 556;

Best Local Similarity 100.0%; Pred. No. 7.8e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDLQERGDNDISPSFGDQPFKD 23

|||||

Db 273 TDLQERGDNDISPSFGDQPFKD 295

RESULT 14

AAU93703

ID AAU93703 standard; peptide; 40 AA.

XX AC AAU93703;

XX DT 02-JUL-2002 (first entry)

XX DE Dental product bone growth enhancing peptide #23.

XX KW Dental product; toothpaste; mouthwash; dental floss; bone growth;
 XX KW integrin binding motif; RGD; skeletal disease; dental disease; tooth;
 XX KW alveolar bone growth; osteoblast; odontoblast; osteopathic.

XX OS Synthetic.

XX PN WO200213775-A1.

XX PD 21-FEB-2002.

XX PF 09-AUG-2001; 2001WO-US25101.

XX PR 16-AUG-2000; 2000US-225879P.

XX PA (BIGB-) BIG BEAR BIO INC.

XX PI Yoneda T, Nomizu M, Kumagai Y;

XX XX WPI; 2002-329525/36.

XX Dental product useful for treating skeletal diseases e.g. dental
 XX diseases comprises a base material and a compound comprising specific
 XX amino acid in a sequence containing the integrin binding motif -

XX Claim 6; Page 20; 44pp; English.

XX The present invention relates to dental products such as toothpastes,
 XX mouthwash and dental floss comprising a base material and a compound
 XX which promotes bone growth. Such compounds are peptide sequences
 XX comprising 10-50 amino acids and containing an integrin binding
 XX motif such as RGD in the D- or L- form, preferably the L-configuration.
 XX The peptides of the invention are useful for treating or preventing
 XX skeletal diseases such as dental disease. The peptides enhance tooth
 XX and/or alveolar bone growth on areas where deterioration has occurred,
 XX as well as the growth or recruiting of osteoblast or odontoblast cells
 XX on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 XX bone growth enhancing peptides for use in dental products.

XX SQ Sequence 40 AA;

Query Match 80.2%; Score 101; DB 23; Length 40;

Best Local Similarity 100.0%; Pred. No. 2.8e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 RGDNDISPSFGDQPFKD 23

|||||

Db 1 RGNNDISPFSGDGQPFKD 18

RESULT 15

AAO20353
ID AAO20353 standard; protein; 40 AA.

XX AC AAO20353;

XX DT 31-MAY-2002 (first entry)

XX DE Protein of matrix extracellular phosphoglycoprotein containing RGD #23.

XX KW Bone growth; RGD motif; integrin binding motif; calcium binding motif;
KW glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
KW alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
KW weakness; matrix extracellular phosphoglycoprotein.

XX OS Unidentified.

XX PN WO200214360-A1.

XX PD 21-FEB-2002.

XX PF 14-AUG-2001; 2001WO-US25542.

XX PR 16-AUG-2000; 2000US-0641034.

XX PR 19-MAR-2001; 2001US-0812485.

XX PA (BIGB-) BIG BEAR BIO INC.

XX PI Kumagai Y, Blacher RW, Yoneda T;

XX DR WPI; 2002-291971/33.

XX PT New peptide compound useful for reducing bone loss, is capable of
PT enhancing bone growth, and comprises an integrin binding motif,
PT glycosaminoglycan binding motif or a calcium binding motif -

XX PS Disclosure; Page 12; 50pp; English.

XX CC The invention relates to a peptide compound capable of enhancing bone
CC growth, and comprising 10-50 amino acids in a sequence, where the amino
CC acids are in D- or L- conformation and the sequence comprises a motif
CC selected from an integrin binding motif, a glycosaminoglycan binding
CC motif and a calcium binding motif. The peptide of the invention is useful
CC for reducing bone loss and for reducing renal phosphate excretion in an
CC individual. The peptide is useful for promoting regeneration of alveolar
CC bone and/or teeth, and increases the number and activity of odontoblasts
CC /osteoclasts that help form dental tissues. The peptide is also useful
CC for treating or preventing a condition associated with skeletal loss or
CC weakness. This sequence represents a protein of a matrix extracellular
CC phosphoglycoprotein containing an RGD motif of the invention.

XX SQ Sequence 40 AA;

Query Match 80.2%; Score 101; DB 23; Length 40;

Best Local Similarity 100.0%; Pred. No. 2.8e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RGNNDISPFSGDGQPFKD 23

|||||

Db 1 RGNNDISPFSGDGQPFKD 18

Search completed: January 29, 2004, 12:38:35
Job time : 43.9737 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:37:19 ; Search time 17.5526 Seconds
(without alignments)
55.442 Million cell updates/sec

Title: US-09-812-485A-49
Perfect score: 126
Sequence: 1 TDLQERGDNDISPFSGDQPFKD 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	41.3	301	2	US-08-355-844-2
2	52	41.3	301	5	PCT-US95-16126-2
3	47	37.3	815	4	US-09-134-001C-3614
4	47	37.3	1198	4	US-09-199-637A-405
5	46	36.5	425	4	US-09-252-991A-31834
6	45.5	36.1	96	2	US-08-341-843B-9
7	45.5	36.1	96	2	US-08-427-497E-14
8	45.5	36.1	101	4	US-09-252-991A-22899
9	45.5	36.1	1260	4	US-08-506-296B-21
10	45	35.7	547	4	US-09-252-991A-32044
11	44	34.9	186	4	US-09-328-352-7703
12	44	34.9	348	1	US-08-176-126B-2
13	44	34.9	348	2	US-08-669-435-2
14	44	34.9	348	5	PCT-US94-14431A-2
15	44	34.9	475	1	US-08-278-091-6
16	44	34.9	475	1	US-08-483-859-6
17	44	34.9	475	1	US-08-472-173-6
18	44	34.9	475	1	US-08-350-741-2
19	44	34.9	475	2	US-08-487-167-6
20	44	34.9	475	2	US-08-482-816-6
21	44	34.9	475	2	US-08-296-149-6
22	44	34.9	475	2	US-08-801-499-6
23	44	34.9	475	2	US-08-463-875A-2
24	44	34.9	475	2	US-08-615-271-6
25	44	34.9	475	3	US-09-074-660-6
26	44	34.9	475	3	US-09-074-659-6
27	44	34.9	475	3	US-09-106-468-6

28	44	34.9	475	3	US-09-106-466A-6	Sequence 6, Appli
29	44	34.9	475	3	US-09-106-467-6	Sequence 6, Appli
30	44	34.9	683	4	US-09-252-991A-19718	Sequence 19718, A
31	43	34.1	169	3	US-08-476-509B-28	Sequence 28, Appli
32	42.5	33.7	306	1	US-08-330-978-1	Sequence 1, Appli
33	42.5	33.7	306	1	US-08-474-042-1	Sequence 1, Appli
34	42.5	33.7	306	1	US-08-484-558-1	Sequence 1, Appli
35	42.5	33.7	306	1	US-08-774-592-1	Sequence 1, Appli
36	42.5	33.7	448	1	US-08-295-411-3	Sequence 3, Appli
37	42.5	33.7	448	2	US-08-955-471-3	Sequence 3, Appli
38	42.5	33.7	448	5	PCT-US92-10068-1	Sequence 3, Appli
39	42.5	33.7	448	5	PCT-US92-10242-3	Sequence 3, Appli
40	42.5	33.7	488	1	US-08-487-037-1	Sequence 1, Appli
41	42.5	33.7	488	4	US-09-367-777-44	Sequence 44, Appli
42	42.5	33.7	488	4	US-09-367-791A-27	Sequence 27, Appli
43	42.5	33.7	545	2	US-08-990-114-1	Sequence 1, Appli
44	42.5	33.7	545	4	US-09-241-333-1	Sequence 1, Appli
45	42	33.3	766	4	US-09-252-991A-31826	Sequence 31826, A

ALIGNMENTS

RESULT 1
US-08-355-844-2
; Sequence 2, Application US/08355844
; Patent No. 5940307
; GENERAL INFORMATION:
; APPLICANT: Fischbarg, Jorge
; APPLICANT: Czegledy, Ferenc
; APPLICANT: Isorovich, Pavel
; APPLICANT: Li, Jun
; APPLICANT: Cheung, Min
; TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
; TITLE OF INVENTION: STRUCTURE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/355,844
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Tang, Henry Y.S.
; REGISTRATION NUMBER: 29,705
; REFERENCE/DOCKET NUMBER: A29927-50/29910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2586
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Rhodobacter capsulatus
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..301
; OTHER INFORMATION: Porin protein
US-08-355-844-2

Query Match 41.3%; Score 52; DB 2; Length 301;
 Best Local Similarity 58.8%; Pred. No. 3.6;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TDLQERGNDSIPSGD 17
 Db 92 TDLDRGNDIPYLTGD 108

RESULT 2

PCT-US95-16126-2
 ; Sequence 2, Application PC/TUS9516126
 ; GENERAL INFORMATION:
 ; APPLICANT: Fischberg, Jorge
 ; APPLICANT: Czegledy, Ferenc
 ; APPLICANT: Iserovich, Pavel
 ; APPLICANT: Li, Jun
 ; APPLICANT: Cheung, Min
 ; TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
 ; STREET: 30 Rockefeller Plaza
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10112-0228
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/16126
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/355,844
 ; FILING DATE: 14-DEC-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Tang, Henry Y.S.
 ; REGISTRATION NUMBER: 29,705
 ; REFERENCE/DOCKET NUMBER: A29927-50/29910
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-408-2586
 ; TELEFAX: 212-765-2519
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 301 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: Rhodobacter capsulatus
 ; FEATURE:
 ; NAME/KEY: Peptide
 ; LOCATION: 1..301
 ; OTHER INFORMATION: Porin protein
 PCT-US95-16126-2

Query Match 41.3%; Score 52; DB 5; Length 301;
 Best Local Similarity 58.8%; Pred. No. 3.6;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TDLQERGNDSIPSGD 17
 Db 92 TDLDRGNDIPYLTGD 108

RESULT 3

US-09-134-001C-3614
 ; Sequence 3614, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 3614
 ; LENGTH: 815
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3614

Query Match 37.3%; Score 47; DB 4; Length 815;
 Best Local Similarity 58.8%; Pred. No. 64;
 Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 7 GNDISPFSGDGPFFKD 23
 Db 481 GDEDQLPSVGPQGVFKD 497

RESULT 4

US-09-199-637A-405
 ; Sequence 405, Application US/09199637A
 ; Patent No. 6355411
 ; GENERAL INFORMATION:
 ; APPLICANT: Ausubel, Frederick
 ; APPLICANT: Goodman, Howard M.
 ; APPLICANT: Rahme, Laurence G.
 ; APPLICANT: Mahajan-Miklos, Shalina
 ; APPLICANT: Tan, Man-Wah
 ; APPLICANT: Cao, Hui
 ; APPLICANT: Drenkard, Eliana
 ; APPLICANT: Tsongalis, John
 ; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
 ; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
 ; FILE REFERENCE: 00786/361002
 ; CURRENT APPLICATION NUMBER: US/09/199,637A
 ; CURRENT FILING DATE: 1998-11-25
 ; PRIOR APPLICATION NUMBER: 60/066,517
 ; PRIOR FILING DATE: 1997-11-25
 ; NUMBER OF SEQ ID NOS: 437
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 405
 ; LENGTH: 1198
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-199-637A-405

Query Match 37.3%; Score 47; DB 4; Length 1198;
 Best Local Similarity 52.4%; Pred. No. 99;
 Matches 11; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 1 TDLQERGNDSIPF---SGD 17
 Db 400 TDLRPSIDNLDLKFPMAYKSGD 420

RESULT 5

US-09-252-991A-31834
 ; Sequence 31834, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.

```

: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
:
:   PRIOR APPLICATION NUMBER: 07/904,991
:
:   FILING DATE: June 26, 1992
:
:   ATTORNEY/AGENT INFORMATION:
:
:     NAME: Minnich, Richard J.
:
:     REGISTRATION NUMBER: 24,175
:
:     REFERENCE/DOCKET NUMBER: CWR 2 149-1
:
: TELECOMMUNICATION INFORMATION:
:
:   TELEPHONE: (216) 861-5582
:
:   TELEFAX: (216) 241-1666
:
:   TELEX: (216) 980162
:
: INFORMATION FOR SEQ ID NO: 9:
:
:   SEQUENCE CHARACTERISTICS:
:
:     LENGTH: 96
:
:     TYPE: amino acid
:
:     STRANDEDNESS: single
:
:     TOPOLOGY: linear
:

```

```

US-08-427-497E-14
; Sequence 14, Application US/08427497E
; Patent No. 5969124
; GENERAL INFORMATION:
; APPLICANT: Lemmon, Vance
; TITLE OF INVENTION: A Method for Characterizing the
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
; Patent No. 5969124
; TITLE OF INVENTION: the Nucleotide Sequence
; TITLE OF INVENTION: Characterized Thereby
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Fay, Sharpe, Beall, Fagan,
; ADDRESSES: Minnich & McKee
; STREET: 1100 Superior Avenue
; STREET: Suite 700
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; MEDIUM TYPE: storable
; COMPUTER: Compaq Prolinea 5100e
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,497E
; FILING DATE: April 24, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/804,991
; FILING DATE: June 26, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Minnich, Richard J.
; REGISTRATION NUMBER: 24,175

```

REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861-5582
TELEFAX: (216) 241-1666
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 96
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acids
HYPOTHETICAL: irrelevant
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: mouse
INDIVIDUAL ISOLATE: 8 day old mouse brain
IMMEDIATE SOURCE:
LIBRARY: lamda GT 10 and lamda GT11
CLONE: synthesis of several clones
PUBLICATION INFORMATION:
AUTHORS: Moos, M.
AUTHORS: Tacke, R.
AUTHORS: Scherer, H.
AUTHORS: Teplow, D.
AUTHORS: Fruh, K.
AUTHORS: Schachner, M.
TITLE: Neural adhesion molecule l1 is a
TITLE: member of the immunoglobulin
TITLE: superfamily with binding domains
TITLE: similar to fibronectin
JOURNAL: NATURE
VOLUME: 334
ISSUE:
PAGES: 701-703
DATE: 1988
US-08-427-497E-14

Query Match 36.1%; Score 45.5; DB 2; Length 96;
Best Local Similarity 61.1%; Pred. No. 9.4;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 DLQERGDNDISPFSGDQ 19
Db 41 DLQERGDSD-KYFIEDCK 57

RESULT 8
US-09-252-991A-22899
Sequence 22899, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22899
LENGTH: 101
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22899

Query Match 36.1%; Score 45.5; DB 4; Length 101;
Best Local Similarity 42.9%; Pred. No. 9.9;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

QY 8 DNDISPFSGD-----GQPF 21
Db 17 DGELKPFSGDTDFIYVGRPF 37

RESULT 9
US-08-506-296B-21
Sequence 21, Application US/08506296B
Patent No. 6313265
GENERAL INFORMATION:
APPLICANT: Phillips, Greg
APPLICANT: Cunningham, Bruce A.
APPLICANT: Crossin, Kathryn L.
TITLE OF INVENTION: NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES
TITLE OF INVENTION: CONTAINING FIBRONECTIN TYPE III REPEATS AND METHODS OF USE
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute
STREET: 10550 No. 6313265th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: California
COUNTRY: U.S.
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/506,296B
FILING DATE: 24-JUL-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI 488.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1260 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-506-296B-21

Query Match 36.1%; Score 45.5; DB 4; Length 1260;
Best Local Similarity 61.1%; Pred. No. 1.8e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 DLQERGDNDISPFSGDQ 19
Db 558 DLQERGDSD-KYFIEDGK 574

RESULT 10
US-09-252-991A-32044
Sequence 32044, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32044


```
;
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32044

Query Match      35.7%; Score 45; DB 4; Length 547;
Best Local Similarity 38.5%; Pred. No. 81;
Matches 10; Conservative 5; Mismatches 7; Indels 4; Gaps 1;

QY 2 DLQERGDNDISPFSGD---GQPFKD 23
Db 411 EADEVGDDDLVFGAGNRQELGQPLHD 436

RESULT 11
US-09-328-352-7703
; Sequence 7703, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: CTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7703
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7703

Query Match      34.9%; Score 44; DB 4; Length 186;
Best Local Similarity 38.1%; Pred. No. 33;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 LQERGDNDISPFSGDQPFKD 23
Db 73 LETHGETDVSPFSGVELYVHD 93

RESULT 12
US-08-176-126B-2
; Sequence 2, Application US/08176126B
; Patent No. 5589358
; GENERAL INFORMATION:
; APPLICANT: Dawson, Paul A.
; TITLE OF INVENTION: IDEAL BILE ACID TRANSPORTER COMPOSITIONS AND
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: US
; ZIP: 77210
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,126B
; FILING DATE: 29-DEC-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: WAKE:002/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577

Query Match      34.9%; Score 44; DB 2; Length 348;
Best Local Similarity 53.3%; Pred. No. 68;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFPS 15
Db 320 TELQEKTDNEMEPRS 334

RESULT 13
US-08-669-435-2
; Sequence 2, Application US/08669435
; Patent No. 5869265
; GENERAL INFORMATION:
; APPLICANT: Dawson, Paul A.
; TITLE OF INVENTION: IDEAL BILE ACID TRANSPORTER COMPOSITIONS AND
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,435
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,126
; FILING DATE: 29-DEC-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: WAKE:002/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577

Query Match      34.9%; Score 44; DB 2; Length 348;
Best Local Similarity 53.3%; Pred. No. 68;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFPS 15
Db 320 TELQEKTDNEMEPRS 334

RESULT 14
```

```
;
; TELEX: na
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-176-126B-2

Query Match      34.9%; Score 44; DB 1; Length 348;
Best Local Similarity 53.3%; Pred. No. 68;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFPS 15
Db 320 TELQEKTDNEMEPRS 334

RESULT 13
US-08-669-435-2
; Sequence 2, Application US/08669435
; Patent No. 5869265
; GENERAL INFORMATION:
; APPLICANT: Dawson, Paul A.
; TITLE OF INVENTION: IDEAL BILE ACID TRANSPORTER COMPOSITIONS AND
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,435
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,126
; FILING DATE: 29-DEC-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: WAKE:002/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577

Query Match      34.9%; Score 44; DB 2; Length 348;
Best Local Similarity 53.3%; Pred. No. 68;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFPS 15
Db 320 TELQEKTDNEMEPRS 334

RESULT 14
```

PCT-US94-14431A-2
; Sequence 2, Application PC/TUS9414431A
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IDEAL BILE ACID TRANSPORTER COMPOSITIONS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: Patent In Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14431A
; FILING DATE: 29-DEC-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/176,126
; FILING DATE: 29-DEC-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: WAKE005P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924 (1) GENERAL INFORMATION:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-14431A-2

Query Match 34.9%; Score 44; DB 5; Length 348;
Best Local Similarity 53.3%; Pred. No. 68;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPPS 15
|:|:|:|:|:|:|:
Db 320 TELQKTDNEMEPSS 334

RESULT 15
US-08-278-091-6
; Sequence 6, Application US/08278091
; Patent No. 5506139
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,091
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-278-091-6

Query Match 34.9%; Score 44; DB 1; Length 475;
Best Local Similarity 57.1%; Pred. No. 97;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 10 DISPFSGDGPQPKD 23
|:|:|:|:|:|:|:
Db 79 DDPFCQDGGSPFQ 92

Search completed: January 29, 2004, 12:41:37
Job time : 17.5526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:38:39 ; Search time 32.6842 Seconds
(without alignments)
146.247 Million cell updates/sec

Title: US-09-812-485A-49
Perfect score: 126
Sequence: 1 TDQERGDNDISPFSGQGPFKD 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	100.0	23	10 US-09-812-485A-49	Sequence 49, Appl
2	126	100.0	23	12 US-10-360-202-1	Sequence 1, Appl
3	126	100.0	24	12 US-10-360-202-17	Sequence 17, Appl
4	126	100.0	24	12 US-10-360-202-18	Sequence 18, Appl
5	126	100.0	97	10 US-09-812-485A-1	Sequence 1, Appl
6	126	100.0	509	12 US-09-794-422-34	Sequence 34, Appl
7	126	100.0	525	10 US-09-814-550-2	Sequence 2, Appl
8	126	100.0	525	12 US-09-794-422-6	Sequence 6, Appl
9	126	100.0	525	12 US-10-311-840-1	Sequence 1, Appl
10	126	100.0	540	12 US-09-794-422-46	Sequence 46, Appl
11	126	100.0	556	12 US-09-794-422-8	Sequence 8, Appl
12	123	97.5	23	12 US-10-360-202-4	Sequence 4, Appl
13	115	91.3	21	12 US-10-360-202-6	Sequence 6, Appl
14	112	88.9	23	12 US-10-360-202-2	Sequence 2, Appl
15	110	87.3	20	12 US-10-360-202-8	Sequence 8, Appl

16	107	84.9	22	12 US-10-360-202-11	Sequence 11, Appl
17	106	84.1	19	12 US-10-360-202-9	Sequence 9, Appl
18	105	83.3	23	12 US-10-360-202-3	Sequence 3, Appl
19	102	81.0	19	12 US-10-360-202-7	Sequence 7, Appl
20	101	80.2	21	12 US-10-360-202-12	Sequence 12, Appl
21	101	80.2	40	10 US-09-812-485A-23	Sequence 23, Appl
22	96	76.2	20	12 US-10-360-202-13	Sequence 13, Appl
23	86	68.3	19	10 US-09-814-550-6	Sequence 6, Appl
24	84	66.7	15	10 US-09-812-485A-48	Sequence 48, Appl
25	84	66.7	33	10 US-09-812-485A-30	Sequence 30, Appl
26	82	65.1	15	10 US-09-812-485A-47	Sequence 47, Appl
27	82	65.1	15	12 US-10-360-202-10	Sequence 10, Appl
28	81	64.3	23	12 US-10-360-202-5	Sequence 5, Appl
29	81	64.3	435	12 US-09-794-422-2	Sequence 2, Appl
30	81	64.3	441	12 US-09-794-422-4	Sequence 4, Appl
31	79	62.7	15	12 US-10-360-202-15	Sequence 15, Appl
32	75	59.5	15	10 US-09-812-485A-46	Sequence 46, Appl
33	73	58.3	40	10 US-09-812-485A-24	Sequence 24, Appl
34	72.5	57.5	30	10 US-09-812-485A-26	Sequence 26, Appl
35	72.5	57.5	35	10 US-09-812-485A-25	Sequence 25, Appl
36	71	56.3	15	12 US-10-360-202-16	Sequence 16, Appl
37	68	54.0	15	12 US-10-360-202-14	Sequence 14, Appl
38	66	52.4	12	10 US-09-812-485A-42	Sequence 42, Appl
39	57	45.2	40	10 US-09-812-485A-32	Sequence 32, Appl
40	57	45.2	45	10 US-09-812-485A-27	Sequence 27, Appl
41	53	42.1	1458	14 US-10-054-691-2	Sequence 2, Appl
42	52	41.3	19	10 US-09-814-550-5	Sequence 5, Appl
43	52	41.3	301	10 US-09-905-176-23	Sequence 23, Appl
44	49	38.9	9	10 US-09-814-550-8	Sequence 8, Appl
45	49	38.9	326	10 US-09-738-626-4959	Sequence 4959, Ap

ALIGNMENTS

RESULT 1
US-09-812-485A-49
; Sequence 49, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russel
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
; FILE REFERENCE: BEAR-006CIP
; CURRENT APPLICATION NUMBER: US/09/812, 485A
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D-00006 peptide
; NAME/KEY: AMIDATION
; LOCATION: 15
US-09-812-485A-49

Query Match 100.0%; Score 126; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDQERGDNDISPFSGQGPFKD 23
|||||
Db 1 TDQERGDNDISPFSGQGPFKD 23

RESULT 2
US-10-360-202-1

; Sequence 1, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-1

Query Match 100.0%; Score 126; DB 12; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDLQERGDNDISPFSGDGPFFKD 23
Db 1 TDLQERGDNDISPFSGDGPFFKD 23

RESULT 3
US-10-360-202-17
; Sequence 17, Application US/10360202
; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-17

Query Match 100.0%; Score 126; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDLQERGDNDISPFSGDGPFFKD 23
Db 2 TDLQERGDNDISPFSGDGPFFKD 24

RESULT 4
US-10-360-202-18
; Sequence 18, Application US/10360202

; Publication No. US20030186891A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Toshio
; APPLICANT: Tsugawa, Naoko
; APPLICANT: Nakagawa, Kimie
; APPLICANT: Blacher, Russell W
; APPLICANT: Kumagai, Yoshinari
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: OF VITAMIN D DEFICIENCY
; FILE REFERENCE: BEAR-010
; CURRENT APPLICATION NUMBER: US/10/360,202
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/335,548
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-18

Query Match 100.0%; Score 126; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDLQERGDNDISPFSGDGPFFKD 23
Db 1 TDLQERGDNDISPFSGDGPFFKD 23

RESULT 5
US-09-812-485A-1
; Sequence 1, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari
; APPLICANT: Blacher, Russell
; APPLICANT: Yoneda, Toshiyuki
; TITLE OF INVENTION: Integrin Binding Motif Containing
; TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
; FILE REFERENCE: BEAR-006CIP
; CURRENT APPLICATION NUMBER: US/09/812,485A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 09/641,034
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptidic compound
US-09-812-485A-1

Query Match 100.0%; Score 126; DB 10; Length 97;
Best Local Similarity 100.0%; Pred. No. 8.1e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDLQERGDNDISPFSGDGPFFKD 23
Db 43 TDLQERGDNDISPFSGDGPFFKD 65

RESULT 6
US-09-794-422-34
; Sequence 34, Application US/09794422
; Publication No. US20030166239A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Thomas A.

```

; APPLICANT: De Wet, Jeffrey R.
; APPLICANT: Gowen, Lori C.
; APPLICANT: Hames, Lynn M.
; TITLE OF INVENTION: Mammalian Osteoregulin
; FILE REFERENCE: PC10445
; CURRENT APPLICATION NUMBER: US/09/794,422
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,617
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/234,500
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-422-34

Query Match      100.0%; Score 126; DB 12; Length 509;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TDQERGDNDISPFSGDGGQPPKD 23
DB      226 TDQERGDNDISPFSGDGGQPPKD 248

RESULT 7
US-09-814-550-2
; Sequence 2, Application US/09814550
; Patent No. US20020102641A1
; GENERAL INFORMATION:
; APPLICANT: Schiavi, Susan
; APPLICANT: Madden, Stephen
; APPLICANT: Manavalan, Parthasarathy
; APPLICANT: Levine, Michael
; APPLICANT: Jan de Beur, Suzanne
; TITLE OF INVENTION: ONCOGENIC OSTEOMALACIA-RELATED GENE 1
; FILE REFERENCE: 5014US
; CURRENT APPLICATION NUMBER: US/09/814,550
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,786
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/241,598
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-814-550-2

Query Match      100.0%; Score 126; DB 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TDQERGDNDISPFSGDGGQPPKD 23
DB      242 TDQERGDNDISPFSGDGGQPPKD 264

RESULT 8
US-09-794-422-6
; Sequence 6, Application US/09794422
; Publication No. US20030166239A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Thomas A.
; APPLICANT: De Wet, Jeffrey R.
; APPLICANT: Gowen, Lori C.
; APPLICANT: Hames, Lynn M.
; TITLE OF INVENTION: Mammalian Osteoregulin
; FILE REFERENCE: PC10445
; CURRENT APPLICATION NUMBER: US/09/794,422
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,617
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/234,500
; PRIOR FILING DATE: 2000-09-22

```


; FEATURE:
; OTHER INFORMATION: Synthetic oligopeptide
US-10-360-202-2

Query Match 88.9%; Score 112; DB 12; Length 23;
Best Local Similarity 91.3%; Pred. No. 2.2e-10;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGDGQPFKD 23

Db 1 TDLQERGDNDISPFSGDGQPFKD 23

RESULT 15

US-10-360-202-8

; Sequence 8, Application US/10360202

; Publication No. US20030186891A1

; GENERAL INFORMATION:

; APPLICANT: Okano, Toshio

; APPLICANT: Tsugawa, Naoko

; APPLICANT: Nakagawa, Kimie

; APPLICANT: Blacher, Russell W

; APPLICANT: Kumagai, Yoshinari

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT

; FILE REFERENCE: BEAR-010

; CURRENT APPLICATION NUMBER: US/10/360,202

; CURRENT FILING DATE: 2003-02-07

; PRIOR APPLICATION NUMBER: 60/335,548

; PRIOR FILING DATE: 2002-02-08

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: Synthetic oligopeptide

US-10-360-202-8

Query Match

Best Local Similarity 87.3%; Score 110; DB 12; Length 20;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPFSGDGQPF 21

Db 1 DLQERGDNDISPFSGDGQPF 20

Search completed: January 29, 2004, 12:42:44

Job time : 33.6842 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:36:28 ; Search time 16.3421 Seconds
(without alignments)
135.349 Million cell updates/sec

Title: US-09-812-485A-49
Perfect score: 126
Sequence: 1 TDLQERGDNDISPFSGQPFKD 23
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.3	301	2 S16070	porin - Rhodobacte
2	52	41.3	320	2 JC5727	porin precursor -
3	49.5	39.3	1259	2 S36126	neural cell adhesi
4	48	38.1	277	2 AD3322	lysozyme (EC 3.2.1
5	48	38.1	385	2 D72858	late expression fa
6	48	38.1	385	2 T41811	LEF-3 orf67 - Bomb
7	47	37.3	444	2 B83891	intracellular alka
8	47	37.3	481	2 AH0410	global stress requ
9	47	37.3	825	2 B89944	hypothetical prote
10	46.5	36.9	1106	2 S38783	integrin alpha cha
11	46.5	36.9	1135	2 T61186	alpha-7 integrin -
12	46.5	36.9	1137	2 JC5950	integrin alpha-7 c
13	46	36.5	308	2 AD3154	nucleotidyltransfe
14	46	36.5	365	2 E98133	hypothetical prote
15	45.5	36.1	347	2 A83170	S-adenosylmethioni
16	45.5	36.1	1260	1 S05479	neural cell adhesi
17	45	35.7	158	2 A11365	a probable repress
18	45	35.7	357	2 C84856	probable protein k
19	45	35.7	393	2 S59499	cellulase egli - s
20	45	35.7	472	2 JQ0802	levansucrase (EC 2
21	45	35.7	851	2 T38173	probable phosphati
22	45	35.7	913	2 D82885	multiple banded an
23	44.5	35.3	698	2 D65210	YibH protein precu
24	44.5	35.3	698	2 D91255	hypothetical prote
25	44.5	35.3	698	2 H86095	hypothetical prote
26	44	34.9	348	2 A49876	Na+-dependent bile
27	44	34.9	475	1 S15337	heat shock protein
28	44	34.9	475	2 AC0528	protease DO protei
29	44	34.9	683	2 D83511	flagellar hook-ass

30	43.5	34.5	284	2 C71507	probable L2 riboso
31	43.5	34.5	284	2 A81661	ribosomal protein
32	43.5	34.5	631	2 D38162	cobr protein - Pee
33	43.5	34.5	1053	2 S44250	integrin alpha-5 c
34	43.5	34.5	1115	2 T09403	integrin alpha cha
35	43.5	34.5	1115	2 T09433	integrin alpha cha
36	43	34.1	140	2 G84608	En/Spm-like transp
37	43	34.1	300	2 F90274	hypothetical prote
38	43	34.1	335	2 G72258	conserved hypothet
39	43	34.1	629	2 T39285	probable transmemb
40	43	34.1	684	1 RNAGVS	transcription init
41	43	34.1	684	2 AF2842	RNA polymerase eig
42	43	34.1	684	2 G97619	RNA polymerase sig
43	43	34.1	762	2 E87592	hypothetical prote
44	43	34.1	779	2 G87573	xanthine dehydroge
45	43	34.1	891	2 E96590	hypothetical prote

ALIGNMENTS

RESULT 1

S16070
porin - Rhodobacter capsulatus
C;Species: Rhodobacter capsulatus
C;Date: 21-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 21-Jul-1995
C;Accession: S16070
R;Schiltz, E.; Kreusch, A.; Nestel, U.; Schulz, G.E.
Eur. J. Biochem. 199, 587-594, 1991
A;Title: Primary structure of porin from Rhodobacter capsulatus.
A;Reference number: S16070; MUID:91330909; PMID:1651239
A;Accession: S16070
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-301 <EUR>
A;Note: the source is designated as Rhodobacter capsulatus

Query Match 41.3%; Score 52; DB 2; Length 301;
Best Local Similarity 58.8%; Pred. No. 3.1;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGD 17
|||:|||||:
Db 92 TDLDDRGNDIPYLTGD 108

RESULT 2

JC5727
porin precursor - Rhodobacter capsulatus
C;Species: Rhodobacter capsulatus
C;Date: 09-Dec-1997 #sequence_revision 23-Jan-1998 #text_change 07-May-1999
C;Accession: JC5727; PC4414
R;Frieschmann, M.D.A.; Fatus, F.; Tadros, M.H.
Gene 183, 61-68, 1996
A;Title: Molecular characterization and organization of porin from Rhodobacter capsulatus
A;Reference number: JC5727; MUID:97149280; PMID:8996088
A;Accession: JC5727
A;Molecule type: DNA
A;Residues: 1-320 <TRI>
A;Cross-references: GB:U57653
A;Experimental source: strain 37B4
A;Accession: PC4414
A;Molecule type: protein
A;Residues: 67-89;131-158;191-218;252-277 <TR2>
C;Comment: This protein is the mutant of porin from Rhodobacter capsulatus strain 37B4. The
ion, which are located exclusively on transmembrane strands.
C;Genetics:
A;Gene: porCa
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-320/Product: porin #status predicted <MAR>

Query Match 41.3%; Score 52; DB 2; Length 320;

Best Local Similarity 58.8%; Pred. No. 3.3;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TDLOERGDNDISPFSGD 17
|||:|||||:
Db 112 TDLDRGNDIPYLTGD 128
|||:|||||:

RESULT 3
S36126
neural cell adhesion molecule L1 - rat
N;Alternate names: nerve growth factor-inducible large external glycoprotein; NILE glyco
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
C;Accession: S36126; S17655; A60917; A30326
R;Miura, M.; Kobayashi, M.; Asou, H.; Uyemura, K.
FEBS Lett. 289, 91-95, 1991
A;Title: Molecular cloning of cDNA encoding the rat neural cell adhesion molecule L1. Th
A;Reference number: S17655; MUID:91372414; PMID:1894011
A;Accession: S36126
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1259 <MIU>
A;Cross-references: EMBL:X59149
A;Accession: S17655
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178,1183-1259 <MI2>
A;Cross-references: EMBL:X59149; NID:956740; PIDN:CAA41860.1; PID:956741
R;Prince, J.T.; Milona, N.; Stallcup, W.B.
J. Neurosci. 9, 1825-1834, 1989
A;Title: Characterization of a partial cDNA clone for the NILE glycoprotein and identifi
A;Reference number: A60917; MUID:89257627; PMID:2723751
A;Accession: A60917
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1159-1199, 'G', 1201-1235, 'K', 1237 <PRI>
A;Note: this paper appeared earlier, with printing errors, as reference A30326
R;Prince, J.T.; Milona, N.; Stallcup, W.B.
J. Neurosci. 9, 876-883, 1989
A;Title: Characterization of a partial cDNA clone for the NILE glycoprotein and identifi
A;Reference number: A30326; MUID:89177485; PMID:2466966
A;Contents: annotation
A;Note: this paper was reprinted as reference A60917 to correct the omission of several
C;Comment: This sequence of this surface-accessible glycoprotein differs at only two pos
accessible only after treatment of cells with detergent and is assumed to be cytoplasmic
C;Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; I
C;Keywords: cell adhesion; duplication; glycoprotein; membrane protein
F;531-592/Domain: immunoglobulin homology <IMM>

Query Match 39.3%; Score 49.5; DB 2; Length 1259;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 2 DLQERGDNDISPFSGDQ 19
|||||:|:
Db 558 DLQERGDSD-KYFIEDGQ 574
|||||:|:

RESULT 4
AD3322
lysozyme (EC 3.2.1.17) - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
C;Accession: AD3322
R;DeIvecchio, V.G.; Kapratl, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3322; PMID:11756688
A;Accession: AD3322
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-277 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL51743.1; PID:gl7982481; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0562
A;Map position: 1
C;Keywords: glycosidase; hydrolase

Query Match 38.1%; Score 48; DB 2; Length 277;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 6 RGDNDISPFSGDQPFK 22
|||||:|:
Db 252 RGDADINTFAGDSASWK 268
|||||:|:

RESULT 5
D72858
late expression factor 3 - Autographa californica nuclear polyhedrosis virus
C;Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
A;Note: dsDNA virus
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 24-Nov-1999
C;Accession: D72858; A40677
R;Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A;Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A;Reference number: A72850; MUID:94303173; PMID:8030224
A;Accession: D72858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-385 <AYR>
A;Cross-references: GB:I22858; NID:9510708; PIDN:AAA66697.1; PID:9559136
R;Li, Y.; Passarelli, A.L.; Miller, L.K.
J. Virol. 67, 5260-5268, 1993
A;Title: Identification, sequence, and transcriptional mapping of lef-3, a baculovirus g
A;Reference number: A40677; MUID:93353600; PMID:8350397
A;Accession: A40677
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-385 <LIA>
A;Cross-references: GB:I18873; NID:9349019; PIDN:AAA02964.1; PID:9349020
A;Note: authors translated the codon ACA for residue 92 as Tyr
C;Genetics:
A;Gene: Ac-lef3
C;Keywords: transcription regulation

Query Match 38.1%; Score 48; DB 2; Length 385;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 8 DNDISPFSGDQPFK 23
|||||:|:
Db 61 DNKIQEYGDSDSFKD 76
|||||:|:

RESULT 6
T41811
LEF-3 orf67 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C;Species: Bombyx mori nuclear polyhedrosis virus, BmsNPV
A;Variety: isolate T3
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T41811
R;Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A;Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A;Reference number: Z22020; MUID:99281911; PMID:10355780
A;Accession: T41811
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-385 <RAM>
A;Cross-references: EMBL:I33180; NID:93745835; PIDN:AAC63740.1; PID:g3745893
A;Experimental source: isolate T3

C/Genetics:
A/Note: 1ef-3

Query Match 38.1%; Score 48; DB 2; Length 385;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 8 DNDISPFSGDGPFXD 23
||| : ||| |||
Db 61 DNKIQBYGDSQSPFXD 76

RESULT 7

B83891 intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain C)
C/Species: Bacillus halodurans

C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C/Accession: B83891

R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A/Reference number: A83650; MUID:20512582; PMID:11058132

A/Accession: B83891

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-444 <STO>

A/Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BAB05649.1; GSPDB:GN00

A/Experimental source: strain C-125

C/Genetics:

A/Gene: aprX

Query Match 37.3%; Score 47; DB 2; Length 444;
Best Local Similarity 47.1%; Pred. No. 28;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPSGSG 18
| : | : | : | : |

Db 319 DTTDREDDVAPFSSRG 335

RESULT 8

AH0410 global stress requirement protein GsrA [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002

C/Accession: AH0410

R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AH0410

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-481 <KUR>

A/Cross-references: GB:AL590842; PIDN:CAC92612.1; PID:g15981308; GSPDB:GN00175

C/Genetics:

A/Gene: gsrA

C/Superfamily: Helicobacter serine proteinase

Query Match 37.3%; Score 47; DB 2; Length 481;
Best Local Similarity 62.5%; Pred. No. 30;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 7 GNDISPFSGDGPFXD 22
||| ||| ||| |||

Db 80 GDN--SPFCQDGSFPFX 93

RESULT 9

B83944 hypothetical protein SA1447 [imported] - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C/Accession: B83944

R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A/Reference number: A89758; MUID:21311952; PMID:11418146

A/Accession: B83944

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-825 <KUR>

A/Cross-references: GB:BA000018; PID:g13701417; PIDN:BAB42711.1; GSPDB:GN00149

A/Experimental source: strain N315

C/Genetics:

A/Gene: SA1447

Query Match 37.3%; Score 47; DB 2; Length 825;
Best Local Similarity 58.8%; Pred. No. 55;
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 GNDISPFSGDGPFXD 23
||| ||| ||| |||

Db 476 GDEDQLPSVGPQGVFXD 492

RESULT 10

S38783 integrin alpha chain - rat (fragment)

C/Species: Rattus norvegicus (Norway rat)

C/Date: 22-Jan-1994 #sequence_revision 14-Jul-1995 #text_change 29-Sep-1999

C/Accession: S38783; S23600

R/Kaufman, S.J.

submitted to the EMBL Data Library, March 1992

A/Reference number: S38783

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1106 <KAU>

A/Cross-references: EMBL:X65036; NID:G56392; PIDN:CAA46170.1; PID:G56393

R/Song, W.K.; Wang, W.; Foster, R.F.; Bielese, D.A.; Kaufman, S.J.

J. Cell Biol. 117, 643-657, 1992

A/Title: H36-alpha7 is a novel integrin alpha chain that is developmentally regulated du

A/Reference number: S23600; MUID:92242309; PMID:1315319

A/Accession: S23600

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-116, 'R', 118-349, 'D', 351-804, 'R', 806, 'V', 808-1106 <SON>

A/Cross-references: EMBL:X65036

C/Superfamily: integrin alpha-2b chain

Query Match 36.9%; Score 46.5; DB 2; Length 1106;
Best Local Similarity 47.8%; Pred. No. 91;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 2 DLQERGDNDI---SPFSGDGPFX 21
||| : | : | : | : |

Db 362 DLNQDGFDPDIAGAPFDGDKVP 384

RESULT 11

I61186 alpha-7 integrin - mouse

C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999

C/Accession: I61186

R/Ziobner, B.L.; Vu, M.P.; Waleh, N.; Crawford, J.; Lin, C.S.; Kramer, R.H.

J. Biol. Chem. 268, 26773-26783, 1993

A/Title: Alternative extracellular and cytoplasmic domains of the integrin alpha 7 subun

A/Reference number: A49691; MUID:94075378; PMID:8253814

A/Accession: I61186

A/Status: preliminary; translated from GB/EMBL/DBJ

```

RESULT 13
AD3154
nucleotidyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AD3154
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.;
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AD3154
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-308 <KUR>
A/Cross-references: GB:AE008689; PIDN:AAU45650.1; PID:gl7743374; GSPDB:GN00187
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu4856
A/Map position: linear chromosome

Query Match 36.5%; Score 46; DB 2; Length 308;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 7 GDNDISPFSGDQPF 21
| : ||| ||||
Db 112 GAREISNFDNGQPF 126

RESULT 14
E98133
Hypothetical protein AGR_L_62 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C/Accession: E98133
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: E98133
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-365 <KUR>
A/Cross-references: GB:AE007870; PIDN:AAK88591.1; PID:gl5158304; GSPDB:GN00170
C/Genetics:
A/Gene: AGR_L_62
A/Map position: linear chromosome

Query Match 36.5%; Score 46; DB 2; Length 365;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 7 GDNDISPFSGDQPF 21
| : ||| ||||
Db 169 GAREISNFDNGQPF 183

RESULT 15
AB3170
S-adenosylmethionine-TRNA ribosyltransferase-isomerase (EC 5.4.99.-) queA PA3824 [similar
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: AB3170
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.D.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

```

A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A83170
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-347 <STO>
A:Cross-references: GB:AE004799; GB:AE004091; NID:g9949981; PIDN:AAG07211.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
C:Gene: queA; PA3824
C:Superfamily: Escherichia coli tRNA ribosyltransferase-isomerase
C:Keywords: intramolecular transferase; isomerase

Query Match	36.1%;	Score	45.5;	DB	2;	Length	347;
Best Local Similarity	42.9%;	Pred. No.	36;				
Matches	9;	Conservative	3;	Mismatches	2;	Indels	7;
				Gaps	1;		

Qy 8 DNDISPFSGD-----GQPF 21
| : : | | | |
| : : | | | |

Db 263 DGLKPFSGDITFIYVGRPF 283
| : : | | | |
| : : | | | |

Search completed: January 29, 2004, 12:40:57
Job time : 17.3421 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:35:18 ; Search time 12.7105 Seconds
(without alignments)
85.096 Million cell updates/sec

Title: US-09-812-485A-49

Perfect score: 126

Sequence: 1 TDLQERGDNDISFFSGDQGFVKD 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	52	41.3	301	PORI_RHOCA	P31243 rhodobacter
2	49.5	39.3	1259	CAML_RAT	Q05695 rattus norv
3	48	38.1	385	LEF3 NPVAC	P41453 autographa
4	46.5	36.9	1106	ITAT7 RAT	O63258 rattus norv
5	46.5	36.9	1179	ITAT7 MOUSE	O61738 mus musculu
6	46.5	36.9	1181	ITAT7 HUMAN	Q13683 homo sapien
7	45.5	36.1	347	QUEA_PSEAE	Q9hxb8 pseudomonas
8	45.5	36.1	1015	ITAT4 DROME	Q9v7a4 drosophila
9	45.5	36.1	1260	CAML MOUSE	P11627 mus musculu
10	45	35.7	301	ITAT2 MOUSE	O8vlg3 mus musculu
11	45	35.7	309	ITAT2 HUMAN	O8wv74 homo sapien
12	45	35.7	393	GUNI_USTMA	P54424 ustilago ma
13	45	35.7	472	SACH_BACAM	P21130 bacillus am
14	45	35.7	508	VLI_HPV60	P50821 human papil
15	45	35.7	851	YDBG SCHPO	Q10366 schizosacch
16	44.5	35.3	698	YJBH ECOLI	P31689 escherichia
17	44.5	35.3	1050	ITAT5 XENLA	O60274 xenopus lae
18	44	34.9	348	NTCI_CRIGR	O60414 c ileal sod
19	44	34.9	475	DEGR_SALTY	P26982 salmonella
20	43.5	34.5	284	RL2 CHLMU	Q9pjl7 chlamydia m
21	43.5	34.5	284	RL2 CHLTR	O84530 chlamydia t
22	43.5	34.5	631	COBT_PSEDE	P29934 pseudomonas
23	43.5	34.5	1053	ITAT5 MOUSE	P11688 mus musculu
24	43.5	34.5	1115	ITAT3 DROME	O44386 drosophila
25	43	34.1	209	COEC_SCYCA	P61130 scyllorhinu
26	43	34.1	684	RPSD_AGRTS	P33452 agrobacteri
27	43	34.1	3703	ABF1 HUMAN	Q15911 homo sapien
28	43	34.1	3726	ABF1 MOUSE	O61329 mus musculu
29	42.5	33.7	488	FAL0 HUMAN	P00742 homo sapien
30	42.5	33.7	1000	ITAT5 DROME	Q9wlm8 drosophila
31	42	33.3	119	GVO1_HALN1	O51968 halobacteri
32	42	33.3	120	GVO1_HALN2	O48313 halobacteri
33	42	33.3	224	DCL_LYCES	Q42463 lycopersico

ALIGNMENTS

RESULT 1									
ID	PORI_RHOCA	STANDARD;	PRT;	301 AA.					
AC	P31243;								
DT	01-JUL-1993 (Rel. 26, Created)								P77608 escherichia
DT	01-JUL-1993 (Rel. 26, Last sequence update)								Q95927 homo sapien
DT	28-FEB-2003 (Rel. 41, Last annotation update)								Q8xnh9 clostridium
DE	Porin.								Q928k1 rhizobium m
OS	Rhodobacter capsulatus (Rhodospseudomonas capsulata).								P32031 drosophila
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;								O54239 rhizobium m
OC	Rhodobacteraceae; Rhodobacter.								Q94390 rhizobium m
OX	NCBI_TaxID=1061;								Q99286 homo sapien
RN	[1]								O33259 mycobacteri
RP	SEQUENCE.								P14565 escherichia
RC	STRAIN=DSM 938 / 37b4;								P20313 bacterioph
RX	MEDLINE=91330909; PubMed=1651239;								P33782 escherichia
RA	Schiltz E., Kreusch A., Nestel U., Schulz G.E.;								
RT	"Primary structure of porin from Rhodobacter capsulatus.";								
RL	Eur. J. Biochem. 199;587-594(1991).								
RN	[2]								
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).								
RC	STRAIN=DSM 938 / 37b4;								
RX	MEDLINE=90336791; PubMed=2165921;								
RA	Weiss M.S., Wacker T., Weckesser J., Welte W., Schulz G.E.;								
RT	"The three-dimensional structure of porin from Rhodobacter capsulatus at 3-A resolution.";								
RL	FEBS Lett. 267:268-272(1990).								
RN	[3]								
RP	X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).								
RC	STRAIN=DSM 938 / 37b4;								
RX	MEDLINE=91192174; PubMed=1707373;								
RA	Weiss M.S., Kreusch A., Schiltz E., Nestel U., Welte W.,								
RT	"The structure of porin from Rhodobacter capsulatus at 1.8-A resolution.";								
RL	FEBS Lett. 280:379-382(1991).								
RN	[4]								
RP	X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).								
RX	MEDLINE=93021091; PubMed=1328651;								
RA	Weiss M.S., Schulz G.E.;								
RT	"Structure of porin refined at 1.8-A resolution.";								
RL	J. Mol. Biol. 227:493-509(1992).								
CC	-!- FUNCTION: FORMS CHANNELS THAT ALLOW THE PASSIVE DIFFUSION OF SMALL HYDROPHILIC SOLUTES UP TO AN EXCLUSION LIMIT OF ABOUT 0.6 kDa.								
CC	-!- SUBUNIT: Homotrimer.								
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.								
CC	-!- DOMAIN: THE MONOMER CONSISTS OF A 16-STRANDED ANTIPARALLEL BETA-SHEET SHEET BARREL (171 RESIDUES), THREE SHORT ALPHA-HELICES (18 RESIDUES) AND 13 HYDROGEN-BONDED REVERSE TURNS (26 RESIDUES).								
DR	PIR; S16070; S16070.								
DR	PDB; 2POR; 15-JUL-93.								
DR	PDB; 3POR; 15-JUL-93.								
KW	Outer membrane; Transmembrane; Porin; 3D-structure.								
FT	STRAND 2 14								
FT	STRAND 19 34								
FT	TURN 36 37								

FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1021 1021 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1029 1029 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1106 1106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1179 1179 Missing (in isoform 2).
 FT VARSPLIC 1179 1182 /FTID=VSP_002592.
 FT SEQUENCE 1259 AA; 140934 MW; 0F12A7C4415F3C08 CRC64;

Query Match 39.3%; Score 49.5; DB 1; Length 1259;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 2 DLQERGDNDISPFSGDQG 19
 |||||:|:|:|
 DB 558 DLQERGDSD-KYFIEDGQ 574

RESULT 3
 LEF3 NPVAC STANDARD; PRT; 385 AA.
 AC P41453;

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Late expression factor 3.
 GN LEF-3.
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 OX NCBI_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L1.
 RX MEDLINE=93353600; PubMed=8350397;
 RA Li Y., Passarelli A.D., Miller L.K.;
 RT "Identification, sequence, and transcriptional mapping of lef-3, a
 RT baculovirus gene involved in late and very late gene expression.";
 RL J. Virol. 67:5260-5268(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 RT polyhedrosis virus.";
 RL Virology 202:586-605(1994).
 CC -!- FUNCTION: REQUIRED FOR LATE AND VERY LATE GENE EXPRESSION. LEF-3
 CC COULD BE A SINGLE STRANDED DNA-BINDING PROTEIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L18873; AAA02964.1; -
 DR EMBL; L22858; AAA66697.1; -
 DR PIR; D72858; D72858.
 KW Early protein; Transcription regulation; DNA-binding.
 SQ SEQUENCE 385 AA; 44551 MW; 9A25ECD7BA7FEDF1 CRC64;

Query Match 38.1%; Score 48; DB 1; Length 385;
 Best Local Similarity 56.2%; Pred. No. 5.6;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 8 DNDISPFSGDGGPFYKD 23
 ||||:|:|:|
 DB 61 DNKIQEYTGDSQSFKD 76

RESULT 4
 ITA7 RAT STANDARD; PRT; 1106 AA.
 AC Q63258; Q63026; Q63027;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-7 (H36-alpha7).
 GN ITGA7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X1B).
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=92242309; PubMed=1315119;
 RA Song W.K., Wang W., Foster R.F., Bielser D.A., Kaufman S.J.;
 RT "H36-alpha 7 is a novel integrin alpha chain that is developmentally
 RT regulated during skeletal myogenesis.";
 RL J. Cell Biol. 117:643-657(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA-7X1A AND ALPHA-7X1C).
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=94171924; PubMed=8126096;
 RA Song W.K., Wang W., Sato H., Bielser D.A., Kaufman S.J.;
 RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal
 RT muscle development: alternate forms, conformational change, and
 RT homologies with serine/threonine kinases and tyrosine phosphatases.";
 RL J. Cell Sci. 106:1139-1152(1993).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=96197133; PubMed=8626012;
 RA Martin P.T., Kaufman S.J., Kramer R.H., Sanes J.R.;
 RT "Synaptic integrins in developing, adult, and mutant muscle: selective
 RT association of alpha1, alpha7A, and alpha7B integrins with the
 RT neuromuscular junction.";
 RL Dev. Biol. 174:125-139(1996).
 CC -!- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
 CC ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
 CC DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
 CC OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
 CC SITES OF SECONDARY FIBER FORMATION. INVOLVED IN THE MAINTENANCE OF
 CC THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR ANCHORAGE,
 CC VIABILITY AND FUNCTIONAL INTEGRITY.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
 CC DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Additional isoforms seem to exist;
 CC Name=Alpha-7X1B;
 CC IsoId=Q63258-1; Sequence=Displayed;
 CC Name=Alpha-7X1A;
 CC IsoId=Q63258-2; Sequence=VSP_002734;
 CC Name=Alpha-7X1C;
 CC IsoId=Q63258-3; Sequence=VSP_002735;
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL AND CARDIAC MUSCLES.
 CC EXPRESSED IN REPLICATING MYOBLASTS. IN DIFFERENTIATED MUSCLE
 CC FIBERS LOCALIZES BETWEEN FIBERS AND THE SURROUNDING MATRIX.
 CC ISOFORMS A AND B ARE EXPRESSED AT MYOTENDINOUS AND NEUROMUSCULAR
 CC JUNCTIONS. ISOFORM C IS EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND
 CC AT EXTRASYNAPTIC SITES.
 CC -!- DEVELOPMENTAL STAGE: ISOFORMS ARE DEVELOPMENTALLY REGULATED DURING
 CC THE FORMATION OF SKELETAL MUSCLE. ISOFORMS A AND C ARE INDUCED
 CC UPON TERMINAL MYOGENIC DIFFERENTIATION; ISOFORM B IS PRESENT
 CC EARLIER IN REPLICATING CELLS AND DIMINISHES UPON DIFFERENTIATION.
 CC -!- PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES IN THE EXTRACELLULAR
 CC DOMAIN IN SKELETAL MYOTUBES (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

RX MEDLINE=95238432; PubMed=7721841;
 RA Zolkiewska A., Moss J.;
 RT "Processing of ADP-ribosylated integrin alpha 7 in skeletal muscle
 myotubes.";
 RL J. Biol. Chem. 270:9227-9233(1995).
 CC -!- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
 CC ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
 CC DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
 CC OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
 CC SITES OF SECONDARY FIBER FORMATION. INVOLVED IN THE MAINTENANCE OF
 CC THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR ANCHORAGE,
 CC VIABILITY AND FUNCTIONAL INTEGRITY. MICE CARRYING A ITGA7 NULL
 CC ALLELE ARE VIABLE AND FERTILE, BUT SHOW PROGRESSIVE MUSCULAR
 CC DYSTROPHY STARTING SOON AFTER BIRTH, BUT WITH A DISTINCT
 CC VARIABILITY IN DIFFERENT MUSCLE TYPES.
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=6;
 CC Comment=Additional isoforms seem to exist. There is a
 CC combination of at least four alternatively spliced domains. two
 CC extracellular (X1 and X2) and two cytoplasmic (A and B). A third
 CC potential alternatively spliced cytoplasmic domain (C) doesn't
 CC appear to be expressed. So far detected are isoforms alpha-7X1A,
 CC alpha-7X1B and alpha-7X2B. Experimental confirmation may be
 CC lacking for some isoforms;
 CC Name=Alpha-7X1X2B;
 CC IsoId=Q61738-1; Sequence=Displayed;
 CC Name=Alpha-7X1A;
 CC IsoId=Q61738-2; Sequence=VSP_002732, VSP_002733;
 CC Name=Alpha-7X1B;
 CC IsoId=Q61738-3; Sequence=VSP_002732;
 CC Name=Alpha-7X2A;
 CC IsoId=Q61738-4; Sequence=VSP_002731, VSP_002733;
 CC Name=Alpha-7X2B;
 CC IsoId=Q61738-5; Sequence=VSP_002731;
 CC Name=Alpha-7X1X2A;
 CC IsoId=Q61738-6; Sequence=VSP_002733;
 CC -!- TISSUE SPECIFICITY: ISOFORMS CONTAINING SEGMENT X2 ARE FOUND IN
 CC ADULT HEART, LUNG AND SKELETAL MUSCLE. ISOFORMS CONTAINING SEGMENT
 CC X1 ARE EXPRESSED IN ADULT HEART, LUNG AND IN PROLIFERATING
 CC SKELETAL MYOBLASTS BUT NOT IN ADULT SKELETAL MUSCLE. ISOFORMS
 CC CONTAINING SEGMENT A ARE EXCLUSIVELY FOUND IN SKELETAL MUSCLE.
 CC ISOFORMS CONTAINING SEGMENT B ARE WIDELY EXPRESSED IN MUSCLE.
 CC FIBERS ISOFORMS CONTAINING SEGMENT A AND B ARE EXPRESSED AT
 CC MYOTENDINOUS AND NEUROMUSCULAR JUNCTIONS; ISOFORMS CONTAINING
 CC SEGMENT C ARE EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND AT
 CC EXTRASYNAPTIC SITES.
 CC -!- DEVELOPMENTAL STAGE: ISOFORMS ARE DEVELOPMENTALLY REGULATED DURING
 CC FORMATION OF SKELETAL MUSCLE. UNDIFFERENTIATED (REPLICATING)
 CC MYOBLASTS EXPRESS ISOFORMS CONTAINING SEGMENT B ONLY, WHEREAS
 CC DIFFERENTIATED MYOBLASTS EXPRESS ISOFORMS CONTAINING SEGMENTS A OR
 CC B.
 CC -!- PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES OF THE EXTRACELLULAR
 CC DOMAIN IN SKELETAL MYOTUBES (IN VITRO).
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L23423; AA16600.1; -;
 CC EMBL; U60419; AAC2772.1; -;
 CC EMBL; Y12380; CAA73023.1; -;
 CC EMBL; Y12383; CAA73023.1; JOINED.
 CC EMBL; Y12384; CAA73023.1; JOINED.

DR EMBL; Y12385; CAA73023.1; JOINED.
 DR EMBL; Y12386; CAA73023.1; JOINED.
 DR EMBL; Y12387; CAA73023.1; JOINED.
 DR EMBL; Y12388; CAA73023.1; JOINED.
 DR EMBL; Y12389; CAA73023.1; JOINED.
 DR EMBL; Y12390; CAA73023.1; JOINED.
 DR EMBL; Y12391; CAA73023.1; JOINED.
 DR EMBL; Y12392; CAA73023.1; JOINED.
 DR EMBL; Y12393; CAA73024.1; -;
 DR EMBL; Y12394; CAA73024.1; -;
 DR EMBL; Y12395; CAA73024.1; JOINED.
 DR EMBL; Y12396; CAA73024.1; JOINED.
 DR EMBL; Y12397; CAA73024.1; JOINED.
 DR EMBL; Y12398; CAA73024.1; JOINED.
 DR EMBL; Y12399; CAA73024.1; JOINED.
 DR EMBL; Y12400; CAA73024.1; JOINED.
 DR EMBL; Y12401; CAA73024.1; JOINED.
 DR EMBL; Y12402; CAA73024.1; JOINED.
 DR EMBL; Y12403; CAA73024.1; JOINED.
 DR EMBL; Y12404; CAA73024.1; JOINED.
 DR EMBL; Y12405; CAA73024.1; JOINED.
 DR EMBL; Y12406; CAA73024.1; JOINED.
 DR EMBL; Y12407; CAA73024.1; JOINED.
 DR EMBL; Y12408; CAA73024.1; JOINED.
 DR EMBL; Y12409; CAA73024.1; JOINED.
 DR EMBL; Y12410; CAA73024.1; JOINED.
 DR EMBL; Y12411; CAA73024.1; JOINED.
 DR EMBL; Y12412; CAA73024.1; JOINED.
 DR EMBL; Y12413; CAA73024.1; JOINED.
 DR EMBL; Y12414; CAA73024.1; JOINED.
 DR EMBL; Y12415; CAA73024.1; JOINED.
 DR EMBL; Y12416; CAA73024.1; JOINED.
 DR EMBL; Y12417; CAA73024.1; JOINED.
 DR EMBL; Y12418; CAA73024.1; JOINED.
 DR EMBL; Y12419; CAA73024.1; JOINED.
 DR EMBL; Y12420; CAA73024.1; JOINED.
 DR EMBL; Y12421; CAA73024.1; JOINED.
 DR EMBL; Y12422; CAA73024.1; JOINED.
 DR EMBL; Y12423; CAA73024.1; JOINED.
 DR EMBL; Y12424; CAA73024.1; JOINED.
 DR EMBL; Y12425; CAA73024.1; JOINED.
 DR EMBL; Y12426; CAA73024.1; JOINED.
 DR EMBL; Y12427; CAA73024.1; JOINED.
 DR EMBL; Y12428; CAA73024.1; JOINED.
 DR EMBL; Y12429; CAA73024.1; JOINED.
 DR EMBL; Y12430; CAA73024.1; JOINED.
 DR EMBL; Y12431; CAA73024.1; JOINED.
 DR EMBL; Y12432; CAA73024.1; JOINED.
 DR EMBL; Y12433; CAA73024.1; JOINED.
 DR EMBL; Y12434; CAA73024.1; JOINED.
 DR EMBL; Y12435; CAA73024.1; JOINED.
 DR EMBL; Y12436; CAA73024.1; JOINED.
 DR EMBL; Y12437; CAA73024.1; JOINED.
 DR EMBL; Y12438; CAA73024.1; JOINED.
 DR EMBL; Y12439; CAA73024.1; JOINED.
 DR EMBL; Y12440; CAA73024.1; JOINED.
 DR EMBL; Y12441; CAA73024.1; JOINED.
 DR EMBL; Y12442; CAA73024.1; JOINED.
 DR EMBL; Y12443; CAA73024.1; JOINED.
 DR EMBL; Y12444; CAA73024.1; JOINED.
 DR EMBL; Y12445; CAA73024.1; JOINED.
 DR EMBL; Y12446; CAA73024.1; JOINED.
 DR EMBL; Y12447; CAA73024.1; JOINED.
 DR EMBL; Y12448; CAA73024.1; JOINED.
 DR EMBL; Y12449; CAA73024.1; JOINED.
 DR EMBL; Y12450; CAA73024.1; JOINED.
 DR EMBL; Y12451; CAA73024.1; JOINED.
 DR EMBL; Y12452; CAA73024.1; JOINED.
 DR EMBL; Y12453; CAA73024.1; JOINED.
 DR EMBL; Y12454; CAA73024.1; JOINED.
 DR EMBL; Y12455; CAA73024.1; JOINED.
 DR EMBL; Y12456; CAA73024.1; JOINED.
 DR EMBL; Y12457; CAA73024.1; JOINED.
 DR EMBL; Y12458; CAA73024.1; JOINED.
 DR EMBL; Y12459; CAA73024.1; JOINED.
 DR EMBL; Y12460; CAA73024.1; JOINED.
 DR EMBL; Y12461; CAA73024.1; JOINED.
 DR EMBL; Y12462; CAA73024.1; JOINED.
 DR EMBL; Y12463; CAA73024.1; JOINED.
 DR EMBL; Y12464; CAA73024.1; JOINED.
 DR EMBL; Y12465; CAA73024.1; JOINED.
 DR EMBL; Y12466; CAA73024.1; JOINED.
 DR EMBL; Y12467; CAA73024.1; JOINED.
 DR EMBL; Y12468; CAA73024.1; JOINED.
 DR EMBL; Y12469; CAA73024.1; JOINED.
 DR EMBL; Y12470; CAA73024.1; JOINED.
 DR EMBL; Y12471; CAA73024.1; JOINED.
 DR EMBL; Y12472; CAA73024.1; JOINED.
 DR EMBL; Y12473; CAA73024.1; JOINED.
 DR EMBL; Y12474; CAA73024.1; JOINED.
 DR EMBL; Y12475; CAA73024.1; JOINED.
 DR EMBL; Y12476; CAA73024.1; JOINED.
 DR EMBL; Y12477; CAA73024.1; JOINED.
 DR EMBL; Y12478; CAA73024.1; JOINED.
 DR EMBL; Y12479; CAA73024.1; JOINED.
 DR EMBL; Y12480; CAA73024.1; JOINED.
 DR EMBL; Y12481; CAA73024.1; JOINED.
 DR EMBL; Y12482; CAA73024.1; JOINED.
 DR EMBL; Y12483; CAA73024.1; JOINED.
 DR EMBL; Y12484; CAA73024.1; JOINED.
 DR EMBL; Y12485; CAA73024.1; JOINED.
 DR EMBL; Y12486; CAA73024.1; JOINED.
 DR EMBL; Y12487; CAA73024.1; JOINED.
 DR EMBL; Y12488; CAA73024.1; JOINED.
 DR EMBL; Y12489; CAA73024.1; JOINED.
 DR EMBL; Y12490; CAA73024.1; JOINED.
 DR EMBL; Y12491; CAA73024.1; JOINED.
 DR EMBL; Y12492; CAA73024.1; JOINED.
 DR EMBL; Y12493; CAA73024.1; JOINED.
 DR EMBL; Y12494; CAA73024.1; JOINED.
 DR EMBL; Y12495; CAA73024.1; JOINED.
 DR EMBL; Y12496; CAA73024.1; JOINED.
 DR EMBL; Y12497; CAA73024.1; JOINED.
 DR EMBL; Y12498; CAA73024.1; JOINED.
 DR EMBL; Y12499; CAA73024.1; JOINED.
 DR EMBL; Y12500; CAA73024.1; JOINED.
 DR EMBL; Y12501; CAA73024.1; JOINED.
 DR EMBL; Y12502; CAA73024.1; JOINED.
 DR EMBL; Y12503; CAA73024.1; JOINED.
 DR EMBL; Y12504; CAA73024.1; JOINED.
 DR EMBL; Y12505; CAA73024.1; JOINED.
 DR EMBL; Y12506; CAA73024.1; JOINED.
 DR EMBL; Y12507; CAA73024.1; JOINED.
 DR EMBL; Y12508; CAA73024.1; JOINED.
 DR EMBL; Y12509; CAA73024.1; JOINED.
 DR EMBL; Y12510; CAA73024.1; JOINED.
 DR EMBL; Y12511; CAA73024.1; JOINED.
 DR EMBL; Y12512; CAA73024.1; JOINED.
 DR EMBL; Y12513; CAA73024.1; JOINED.
 DR EMBL; Y12514; CAA73024.1; JOINED.
 DR EMBL; Y12515; CAA73024.1; JOINED.
 DR EMBL; Y12516; CAA73024.1; JOINED.
 DR EMBL; Y12517; CAA73024.1; JOINED.
 DR EMBL; Y12518; CAA73024.1; JOINED.
 DR EMBL; Y12519; CAA73024.1; JOINED.
 DR EMBL; Y12520; CAA73024.1; JOINED.
 DR EMBL; Y12521; CAA73024.1; JOINED.
 DR EMBL; Y12522; CAA73024.1; JOINED.
 DR EMBL; Y12523; CAA73024.1; JOINED.
 DR EMBL; Y12524; CAA73024.1; JOINED.
 DR EMBL; Y12525; CAA73024.1; JOINED.
 DR EMBL; Y12526; CAA73024.1; JOINED.
 DR EMBL; Y12527; CAA73024.1; JOINED.
 DR EMBL; Y12528; CAA73024.1; JOINED.
 DR EMBL; Y12529; CAA73024.1; JOINED.
 DR EMBL; Y12530; CAA73024.1; JOINED.
 DR EMBL; Y12531; CAA73024.1; JOINED.
 DR EMBL; Y12532; CAA73024.1; JOINED.
 DR EMBL; Y12533; CAA73024.1; JOINED.
 DR EMBL; Y12534; CAA73024.1; JOINED.
 DR EMBL; Y12535; CAA73024.1; JOINED.
 DR EMBL; Y12536; CAA73024.1; JOINED.
 DR EMBL; Y12537; CAA73024.1; JOINED.
 DR EMBL; Y12538; CAA73024.1; JOINED.
 DR EMBL; Y12539; CAA73024.1; JOINED.
 DR EMBL; Y12540; CAA73024.1; JOINED.
 DR EMBL; Y12541; CAA73024.1; JOINED.
 DR EMBL; Y12542; CAA73024.1; JOINED.
 DR EMBL; Y12543; CAA73024.1; JOINED.
 DR EMBL; Y12544; CAA73024.1; JOINED.
 DR EMBL; Y12545; CAA73024.1; JOINED.
 DR EMBL; Y12546; CAA73024.1; JOINED.
 DR EMBL; Y12547; CAA73024.1; JOINED.
 DR EMBL; Y12548; CAA73024.1; JOINED.
 DR EMBL; Y12549; CAA73024.1; JOINED.
 DR EMBL; Y12550; CAA73024.1; JOINED.
 DR EMBL; Y12551; CAA73024.1; JOINED.
 DR EMBL; Y12552; CAA73024.1; JOINED.
 DR EMBL; Y12553; CAA73024.1; JOINED.
 DR EMBL; Y12554; CAA73024.1; JOINED.
 DR EMBL; Y12555; CAA73024.1; JOINED.
 DR EMBL; Y12556; CAA73024.1; JOINED.
 DR EMBL; Y12557; CAA73024.1; JOINED.
 DR EMBL; Y12558; CAA73024.1; JOINED.
 DR EMBL; Y12559; CAA73024.1; JOINED.
 DR EMBL; Y12560; CAA73024.1; JOINED.
 DR EMBL; Y12561; CAA73024.1; JOINED.
 DR EMBL; Y12562; CAA73024.1; JOINED.
 DR EMBL; Y12563; CAA73024.1; JOINED.
 DR EMBL; Y12564; CAA73024.1; JOINED.
 DR EMBL; Y12565; CAA73024.1; JOINED.
 DR EMBL; Y12566; CAA73024.1; JOINED.
 DR EMBL; Y12567; CAA73024.1; JOINED.
 DR EMBL; Y12568; CAA73024.1; JOINED.
 DR EMBL; Y12569; CAA73024.1; JOINED.
 DR EMBL; Y12570; CAA73024.1; JOINED.
 DR EMBL; Y12571; CAA73024.1; JOINED.
 DR EMBL; Y12572; CAA73024.1; JOINED.
 DR EMBL; Y12573; CAA73024.1; JOINED.
 DR EMBL; Y12574; CAA73024.1; JOINED.
 DR EMBL; Y12575; CAA73024.1; JOINED.
 DR EMBL; Y12576; CAA73024.1; JOINED.
 DR EMBL; Y12577; CAA73024.1; JOINED.
 DR EMBL; Y12578; CAA73024.1; JOINED.
 DR EMBL; Y12579; CAA73024.1; JOINED.
 DR EMBL; Y12580; CAA73024.1; JOINED.
 DR EMBL; Y12581; CAA73024.1; JOINED.
 DR EMBL; Y12582; CAA73024.1; JOINED.
 DR EMBL; Y12583; CAA73024.1; JOINED.
 DR EMBL; Y12584; CAA73024.1; JOINED.
 DR EMBL; Y12585; CAA73024.1; JOINED.
 DR EMBL; Y12586; CAA73024.1; JOINED.
 DR EMBL; Y12587; CAA73024.1; JOINED.
 DR EMBL; Y12588; CAA73024.1; JOINED.
 DR EMBL; Y12589; CAA73024.1; JOINED.
 DR EMBL; Y12590; CAA73024.1; JOINED.
 DR EMBL; Y12591; CAA73024.1; JOINED.
 DR EMBL; Y12592; CAA73024.1; JOINED.
 DR EMBL; Y12593; CAA73024.1; JOINED.
 DR EMBL; Y12594; CAA73024.1; JOINED.
 DR EMBL; Y12595; CAA73024.1; JOINED.
 DR EMBL; Y12596; CAA73024.1; JOINED.
 DR EMBL; Y12597; CAA73024.1; JOINED.
 DR EMBL; Y12598; CAA73024.1; JOINED.
 DR EMBL; Y12599; CAA73024.1; JOINED.
 DR EMBL; Y12600; CAA73024.1; JOINED.
 DR EMBL; Y12601; CAA73024.1; JOINED.
 DR EMBL; Y12602; CAA73024.1; JOINED.
 DR EMBL; Y12603; CAA73024.1; JOINED.
 DR EMBL; Y12604; CAA73024.1; JOINED.
 DR EMBL; Y12605; CAA73024.1; JOINED.
 DR EMBL; Y12606; CAA73024.1; JOINED.
 DR EMBL; Y12607; CAA73024.1; JOINED.
 DR EMBL; Y12608; CAA73024.1; JOINED.
 DR EMBL; Y12609; CAA73024.1; JOINED.
 DR EMBL; Y12610; CAA73024.1; JOINED.
 DR EMBL; Y12611; CAA73024.1; JOINED.
 DR EMBL; Y12612; CAA73024.1; JOINED.
 DR EMBL; Y12613; CAA73024.1; JOINED.
 DR EMBL; Y12614; CAA73024.1; JOINED.
 DR EMBL; Y12615; CAA73024.1; JOINED.
 DR EMBL; Y12616; CAA73024.1; JOINED.
 DR EMBL; Y12617; CAA73024.1; JOINED.
 DR EMBL; Y12618; CAA73024.1; JOINED.
 DR EMBL; Y12619; CAA73024.1; JOINED.
 DR EMBL; Y12620; CAA73024.1; JOINED.
 DR EMBL; Y12621; CAA73024.1; JOINED.
 DR EMBL; Y12622; CAA73024.1; JOINED.
 DR EMBL; Y12623; CAA73024.1; JOINED.
 DR EMBL; Y12624; CAA73024.1; JOINED.
 DR EMBL; Y12625; CAA73024.1; JOINED.
 DR EMBL; Y12626; CAA73024.1; JOINED.
 DR EMBL; Y12627; CAA73024.1; JOINED.
 DR EMBL; Y12628; CAA73024.1; JOINED.
 DR EMBL; Y12629; CAA73024.1; JOINED.
 DR EMBL; Y12630; CAA73024.1; JOINED.
 DR EMBL; Y12631; CAA73024.1; JOINED.
 DR EMBL; Y12632; CAA73024.1; JOINED.
 DR EMBL; Y12633; CAA73024.1; JOINED.
 DR EMBL; Y12634; CAA73024.1; JOINED.
 DR EMBL; Y12635; CAA73024.1; JOINED.
 DR EMBL; Y12636; CAA73024.1; JOINED.
 DR EMBL; Y12637; CAA73024.1; JOINED.
 DR EMBL; Y12638; CAA73024.1; JOINED.
 DR EMBL; Y12639; CAA73024.1; JOINED.
 DR EMBL; Y12640; CAA73024.1; JOINED.
 DR EMBL; Y12641; CAA73024.1; JOINED.
 DR EMBL; Y12642; CAA73024.1; JOINED.
 DR EMBL; Y12643; CAA73024.1; JOINED.
 DR EMBL; Y12644; CAA73024.1; JOINED.
 DR EMBL; Y12645; CAA73024.1; JOINED.
 DR EMBL; Y12646; CAA73024.1; JOINED.
 DR EMBL; Y12647; CAA73024.1; JOINED.
 DR EMBL; Y12648; CAA73024.1; JOINED.
 DR EMBL; Y12649; CAA73024.1; JOINED.
 DR EMBL; Y12650; CAA73024.1; JOINED.
 DR EMBL; Y12651; CAA73024.1; JOINED.
 DR EMBL; Y12652; CAA73024.1; JOINED.
 DR EMBL; Y12653; CAA73024.1; JOINED.
 DR EMBL; Y12654; CAA73024.1; JOINED.
 DR EMBL; Y12655; CAA73024.1; JOINED.
 DR EMBL; Y12656; CAA73024.1; JOINED.
 DR EMBL; Y12657; CAA73024.1; JOINED.
 DR EMBL; Y12658; CAA73024.1; JOINED.
 DR EMBL; Y12659; CAA73024.1; JOINED.
 DR EMBL; Y12660; CAA73024.1; JOINED.
 DR EMBL; Y12661; CAA73024.1; JOINED.
 DR EMBL; Y12662; CAA73024.1; JOINED.
 DR EMBL; Y12663; CAA73024.1; JOINED.
 DR EMBL; Y12664; CAA73024.1; JOINED.
 DR EMBL; Y12665; CAA73024.1; JOINED.
 DR EMBL; Y12666; CAA73024.1; JOINED.
 DR EMBL; Y12667; CAA73024.1; JOINED.
 DR EMBL; Y12668; CAA73024.1; JOINED.
 DR EMBL; Y12669; CAA73024.1; JOINED.
 DR EMBL; Y12670; CAA73024.1; JOINED.
 DR EMBL; Y12671; CAA73024.1; JOINED.
 DR EMBL; Y12672; CAA73024.1; JOINED.
 DR EMBL; Y12673; CAA73024.1; JOINED.
 DR EMBL; Y12674; CAA73024.1; JOINED.
 DR EMBL; Y12675; CAA73024.1; JOINED.
 DR EMBL; Y12676; CAA73024.1; JOINED.
 DR EMBL; Y12677; CAA73024.1; JOINED.
 DR EMBL; Y12678; CAA73024.1; JOINED.
 DR EMBL; Y12679; CAA73024.1; JOINED.
 DR EMBL; Y12680; CAA73024.1; JOINED.
 DR EMBL; Y12681; CAA73024.1; JOINED.
 DR EMBL; Y12682; CAA73024.1; JOINED.
 DR EMBL; Y12683; CAA73024.1; JOINED.
 DR EMBL; Y12684; CAA73024.1; JOINED.
 DR EMBL; Y12685; CAA73024.1; JOINED.
 DR EMBL; Y12686; CAA73024.1; JOINED.
 DR EMBL; Y12687; CAA73024.1; JOINED.
 DR EMBL; Y12688; CAA73024.1; JOINED.
 DR EMBL; Y12689; CAA73024.1; JOINED.
 DR EMBL; Y12690; CAA73024.1; JOINED.
 DR EMBL; Y12691; CAA73024.1; JOINED.
 DR EMBL; Y12692; CAA73024.1; JOINED.
 DR EMBL; Y12693; CAA73024.1; JOINED.
 DR EMBL; Y12694; CAA73024.1; JOINED.
 DR EMBL; Y12695; CAA73024.1; JOINED.
 DR EMBL; Y12696; CAA73024.1; JOINED.
 DR EMBL; Y12697; CAA73024.1; JOINED.
 DR EMBL; Y12698; CAA73024.1; JOINED.
 DR EMBL; Y12699; CAA73024.1; JOINED.
 DR EMBL; Y12700; CAA73024.1; JOINED.
 DR EMBL; Y12701; CAA73024.1; JOINED.
 DR EMBL; Y12702; CAA73024.1; JOINED.
 DR EMBL; Y12703; CAA73024.1; JOINED.
 DR EMBL; Y12704; CAA73024.1; JOINED.
 DR EMBL; Y12705; CAA73024.1; JOINED.
 DR EMBL; Y12706; CAA73024.1; JOINED.
 DR EMBL; Y12707; CAA73024.1; JOINED.
 DR EMBL; Y12708; CAA73024.1; JOINED.
 DR EMBL; Y12709; CAA73024.1; JOINED.
 DR EMBL; Y12710; CAA73024.1; JOINED.
 DR EMBL; Y12711; CAA73024.1; JOINED.
 DR EMBL; Y12712; CAA73024.1; JOINED.
 DR EMBL; Y12713; CAA73024.1; JOINED.
 DR EMBL; Y12714; CAA73024.1; JOINED.
 DR EMBL; Y12715; CAA73024.1; JOINED.
 DR EMBL; Y12716; CAA73024.1; JOINED.
 DR EMBL; Y12717; CAA73024.1; JOINED.
 DR EMBL; Y12718; CAA73024.1; JOINED.
 DR EMBL; Y12719; CAA73024.1; JOINED.
 DR EMBL; Y12720; CAA73024.1; JOINED.
 DR EMBL; Y12721; CAA73024.1; JOINED.
 DR EMBL; Y12722; CAA73024.1; JOINED.
 DR EMBL; Y12723; CAA73024.1; JOINED.
 DR EMBL; Y12724; CAA73024.1; JOINED.
 DR EMBL; Y12725; CAA73024.1; JOINED.
 DR EMBL; Y12726; CAA73024.1; JOINED.
 DR EMBL; Y12727; CAA73024.1; JOINED.
 DR EMBL; Y12728; CAA73024.1; JOINED.
 DR EMBL; Y12729; CAA73024.1; JOINED.
 DR EMBL; Y12730; CAA73024.1; JOINED.
 DR EMBL; Y12731; CAA73024.1; JOINED.
 DR EMBL; Y12732; CAA73024.1; JOINED.
 DR EMBL; Y12733; CAA73024.1; JOINED.
 DR EMBL; Y12734; CAA73024.1; JOINED.
 DR EMBL; Y12735; CAA73024.1; JOINED.
 DR EMBL; Y12736; CAA73024.1; JOINED.
 DR EMBL; Y12737; CAA73024.1; JOINED.
 DR EMBL; Y12738; CAA73024.1; JOINED.
 DR EMBL; Y12739; CAA73024.1; JOINED.
 DR EMBL; Y12740; CAA73024.1; JOINED.
 DR EMBL; Y12741; CAA73024.1; JOINED.
 DR EMBL; Y12742; CAA73024.1; JOINED.
 DR EMBL; Y12743; CAA73024.1; JOINED.
 DR EMBL; Y12744; CAA73024.1; JOINED.
 DR EMBL; Y12745; CAA73024.1; JOINED.
 DR EMBL; Y12746; CAA73024.1; JOINED.
 DR EMBL; Y12747; CAA73024.1; JOINED.
 DR EMBL; Y12748; CAA73024.1; JOINED.
 DR EMB

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-7 precursor.
GN ITGA7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA-7X2B AND ALPHA-7X2DB).
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=98139911; PubMed=9473524;
RA Leung E., Lim S.P., Berg R., Yang Y., Ni J., Wang S.-X.,
RA Kissansen G.W.;
RT "A novel extracellular domain variant of the human integrin alpha 7
RT subunit generated by alternative intron splicing.";
RL Biochem. Biophys. Res. Commun. 243:317-325(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).
RX MEDLINE=98250181; PubMed=9590299;
RA Hayashi Y.K., Chou F.-J., Engvall E., Ogawa M., Matsuda C.,
RA Hirabayashi S., Yokochi K., Ziober B.L., Kramer R.H., Kaufman S.J.,
RA Orawa E., Goto Y.-I., Nonaka I., Tsukahara T., Wang J.Z.,
RA Hoffman E.P., Arahata K.;
RT "Mutations in the integrin alpha7 gene cause congenital myopathy.";
RL Nat. Genet. 19:94-97(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).
RA Vazirianakis I.S., Ziober B.L., Kramer R.H.;
RT "Cloning of human integrin alpha-7 cDNA.";
RN Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Skeletal muscle;
RX MEDLINE=99333684; PubMed=10403775;
RA Vignier N., Moghadaszadeh B., Gary F., Beckmann J., Mayer U.,
RA Guicheney P.;
RT "Structure, genetic localization, and identification of the cardiac
RT and skeletal muscle transcripts of the human integrin alpha7 gene
RT (ITGA7).";
RL Biochem. Biophys. Res. Commun. 260:357-364(1999).
RN [5]
RP SEQUENCE OF 34-45.
RC TISSUE=Melanoma;
RX MEDLINE=92198982; PubMed=18393357;
RA Kramer R.H., Vu M.P., Cheng Y.F., Ramos D.M., Timpl R., Waleh N.;
RT "Laminin-binding integrin alpha 7 beta 1: functional characterization
RT and expression in normal and malignant melanocytes.";
RL Cell Regul. 2:805-817(1991).
RN [6]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2).
RC TISSUE=Heart;
RX MEDLINE=94075378; PubMed=8253814;
RA Ziober B.L., Vu M.P., Waleh N., Crawford J., Lin C.-S., Kramer R.H.;
RT "Alternative extracellular and cytoplasmic domains of the integrin
RT alpha 7 subunit are differentially expressed during development.";
RL J. Biol. Chem. 268:26773-26783(1993).
RN [7]
RP SEQUENCE OF 1105-1181 FROM N.A. (ISOFORM ALPHA-7B).
RC TISSUE=Fetal muscle;
RX MEDLINE=94171924; PubMed=8126096;
RA Song W.K., Wang W., Sato H., Bielsner D.A., Kaufman S.J.;
RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal
RT muscle development: alternate forms, conformational change, and
RT homologies with serine/threonine kinases and tyrosine phosphatases.";
RL J. Cell Sci. 106:1139-1152(1993).
RN [8]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7A).
RC TISSUE=Skeletal muscle;
RX MEDLINE=98012902; PubMed=9352853;
RA Basora N., Vachon P.H., Herring-Gilliam F.E., Perreault N.,
RA Beaulieu J.-F.;

RT "Relation between integrin alpha7Bbeta1 expression in human intestinal
RL cells and enterocytic differentiation.";
RN Gastroenterology 113:1510-1521(1997).
RP FUNCTION
RX MEDLINE=97453229; PubMed=9307969;
RA Ziober B.L., Chen Y.Q., Kramer R.H.;
RT "The laminin-binding activity of the alpha 7 integrin receptor is
RT defined by developmentally regulated splicing in the extracellular
RT domain.";
RL Mol. Biol. Cell 8:1723-1734(1997).
RN [10]
RP FUNCTION.
RX MEDLINE=20160722; PubMed=10694445;
RA Schieber S., Mielenz D., Echtermeyer F., Hapke S., Poeschl E.,
RA von der Mark H., Moch H., von der Mark K.;
RT "The role of extracellular and cytoplasmic splice domains of
RT alpha7-integrin in cell adhesion and migration on laminins.";
RN Exp. Cell Res. 255:303-313(2000).
RN [11]
RP TISSUE SPECIFICITY.
RX MEDLINE=96197133; PubMed=8626012;
RA Martin P.F., Kaufman S.J., Kramer R.H., Sanes J.R.;
RT "Synaptic integrins in developing, adult, and mutant muscle: selective
RT association of alpha1, alpha7A, and alpha7B integrins with the
RT neuromuscular junction.";
RL Dev. Biol. 174:125-139(1996).
RN [12]
RP FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
SITES OF SECONDARY FIBER FORMATION. IT IS INVOLVED IN THE
MAINTENANCE OF THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR
ANCHORAGE, VIABILITY AND FUNCTIONAL INTEGRITY. ISOFORMS ALPHA-7X2B
AND ALPHA-7X1B PROMOTE MYOBLAST MIGRATION ON LAMININ 1 AND LAMININ
2/4, BUT ALPHA-7X1B IS LESS ACTIVE ON LAMININ 1 (IN VITRO).
SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
SUBCELLULAR LOCATION: Type I membrane protein.
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=12;
Comment=Additional isoforms seem to exist. There is a
combination of at least five alternatively spliced domains,
three extracellular (X1, X2 and D) and two cytoplasmic (A and
B). A third potential alternatively spliced cytoplasmic domain
(C) does not appear to be expressed. In vitro generated isoform
X2C shows function. So far detected are isoforms Alpha-7X1A,
Alpha-7X2B (major), Alpha-7X2DB (minor) and minor isoforms
containing segment X1X2. Experimental confirmation may be
lacking for some isoforms;
Name=Alpha-7X1X2B;
IsoId=Q13683-1; Sequence=Displayed;
Name=Alpha-7X1A;
IsoId=Q13683-2; Sequence=VSP_002728, VSP_002730;
Name=Alpha-7X1B;
IsoId=Q13683-3; Sequence=VSP_002728;
Name=Alpha-7X1DA;
IsoId=Q13683-4; Sequence=VSP_002728, VSP_002729, VSP_002730;
Name=Alpha-7X1DB;
IsoId=Q13683-5; Sequence=VSP_002728, VSP_002729;
Name=Alpha-7X2A;
IsoId=Q13683-6; Sequence=VSP_002727, VSP_002730;
Name=Alpha-7X2B;
IsoId=Q13683-7; Sequence=VSP_002727;
Name=Alpha-7X2DA;
IsoId=Q13683-8; Sequence=VSP_002727, VSP_002729, VSP_002730;
Name=Alpha-7X2DB;
IsoId=Q13683-9; Sequence=VSP_002727, VSP_002729;
Name=Alpha-7X1X2A;
IsoId=Q13683-10; Sequence=VSP_002730;
Name=Alpha-7X1X2DA;
IsoId=Q13683-11; Sequence=VSP_002729, VSP_002730;


```
FT REPEAT 44 66 MORN 2.
FT REPEAT 67 89 MORN 3.
FT REPEAT 90 112 MORN 4.
FT REPEAT 113 135 MORN 5.
FT REPEAT 159 181 MORN 6.
FT REPEAT 301 AA; 34181 MW; 704ABEFB94DFC90E CRC64;
SQ SEQUENCE 301 AA; 34181 MW; 704ABEFB94DFC90E CRC64;

Query Match 35.7%; Score 45; DB 1; Length 301;
Best Local Similarity 37.5%; Pred. No. 12;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPFSGD 17
Db 8 ELEEGENDLGEYEGE 23

RESULT 11
TS22 HUMAN STANDARD; PRT; 309 AA.
AC Q8WYR4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Testis-specific gene A2 (Male meiotic metaphase chromosome-associated
DE acidic protein) (Meichroacidin).
GN TSGA2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98246622; PubMed=9578619;
RA Tsuchida J., Nishina Y., Wakabayashi N., Nozaki M., Sakai Y.,
RA Nishimune Y.;
RT "Molecular cloning and characterization of meichroacidin (male meiotic
RT metaphase chromosome-associated acidic protein).";
RL Dev. Biol. 197:67-76(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Shimizu N., Kudoh J., Shibuya K.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May play an important role in male meiosis (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic in late spermatocytes, secondary
CC spermatocytes and round spermatids. Gathered around metaphase
CC chromosomes during meiotic divisions (By similarity).
CC -1- TISSUE SPECIFICITY: Testis specific.
CC -1- SIMILARITY: Contains 6 MORN repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
FT REPEAT 90 112 MORN 4.
FT REPEAT 113 135 MORN 5.
FT REPEAT 159 181 MORN 6.
SQ SEQUENCE 309 AA; 35124 MW; 09A20A766D5A6AB8 CRC64;

Query Match 35.7%; Score 45; DB 1; Length 309;
Best Local Similarity 46.7%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPFSG 16
Db 8 ELEEGENDIGYEG 22

RESULT 12
GUN1_USTMA STANDARD; PRT; 393 AA.
AC P5424;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Endoglucanase 1 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 1)
DE (Cellulase 1) (EG 1).
GN EGL1.
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=5270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PB11;
RX MEDLINE=96145728; PubMed=8590631;
RA Schaevecker F., Wanner G., Kahmann R.;
RT Filament-specific expression of a cellulase gene in the dimorphic
RT fungus Ustilago maydis.
RL Biol. Chem. Hoppe-Seyler 376:617-625(1995).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose, lichenin and cereal beta-D-glucans.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: HYPHAL TIP.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN FILAMENTOUS DIKARYON.
CC -1- PTM: MAY ALSO BE O-GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY K (FAMILY 45 OF GLYCOSYL
CC HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```



```
Db 235 QKQDNTISPSYG 247
| : ||| ||| : ||
RESULT 13
SACB_BACAM STANDARD; PRT; 472 AA.
AC P21130;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Levanucrase precursor (EC 2.4.1.10) (Beta-D-fructofuranosyl
transferase) (Sucrose 6-fructosyl transferase).
GN SACB.
OS Bacillus amyloliquefaciens.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1390;
RN (1)_TaxID=1390;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23844;
RX MEDLINE=91092506; PubMed=2265762;
RA Tang L.B., Lenstra R., Borchert T.V., Vasantha N.;
RT "Isolation and characterization of levanucrase-encoding gene from
RT Bacillus amyloliquefaciens.";
RL Gene 96:89-93(1990).
CC glucose + {(2,6)-beta-D-fructosyl}(N) =
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- INDUCTION: By sucrose.
CC -1- SIMILARITY: BELONGS TO FAMILY 68 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52989; CAA37179.1; -
DR PIR; JQ0802; JQ0802.
DR InterPro; IPR003469; Glyco_hydro_68.
DR Pfam; PF02435; Glyco_hydro_68; 1.
DR Transferase; Glycosyltransferase; Signal.
FT SIGNAL 1 29
FT CHAIN 30 472 LEVANSUCRASE.
SQ SEQUENCE 472 AA; 52859 MW; F38592D272677E7D CRC64;

Query Match 35.7%; Score 45; DB 1; Length 472;
Best Local Similarity 45.0%; Pred. No. 21;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 4 QERGDNDISPSFGDGGPFKD 23
| : ||| : |||
DB 126 QKVGDSIDSQKNGRVFKD 145

RESULT 14
VLI_HPV60 STANDARD; PRT; 508 AA.
AC P50821; Q80947;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 60.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=40540;
RN (1)
RP SEQUENCE FROM N.A.
RA Delius H.;

Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 372-467 FROM N.A.
RA Chan S.Y., Delius H., Halpern A.L., Bernard H.U.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U31792; AAA79491.1; -
DR EMBL; U21876; AAA92837.1; -
DR InterPro; IPR002210; PV_capsid_L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPV_CAPSID_L1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; Late protein.
SQ SEQUENCE 508 AA; 57827 MW; 3149549895534D00 CRC64;

Query Match 35.7%; Score 45; DB 1; Length 508;
Best Local Similarity 64.3%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPPFS 15
| ||| | : |||
DB 459 DLQERFNSNELSQFS 472

RESULT 15
YDBG_SCHPO STANDARD; PRT; 851 AA.
AC Q10366;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C22E12.16c in chromosome I.
GN SPAC22E12.16C
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN (1)_TaxID=4896;
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
```

```
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: STRONG TO YEAST PIK1.
CC -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z70043; CAA93903.1; -.
CC PIR; T38173; T38173.
CC GeneDB_Spombe; SPAC22B12.16C; -.
CC InterPro; IPR000403; PI3_PI4_kinase.
CC Pfam; PF00454; PI3_PI4_kinase; 1.
CC SMART; SM00146; PI3KC; 1.
CC PROSITE; PS00915; PI3_4_KINASE_1; 1.
CC PROSITE; PS00916; PI3_4_KINASE_2; 1.
CC PROSITE; PS0290; PI3_4_KINASE_3; 1.
KW Hypothetical protein; Transferase; Kinase.
FT DOMAIN 581 827 PI3K/PI4K.
SQ SEQUENCE 851 AA; 96657 MW; A991E3C7E3D980BE CRC64;

Query Match 35.7%; Score 45; DB 1; Length 851;
Best Local Similarity 50.0%; Pred.No. 41;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPPSGDG 18
Db :||: ||||| :|
382 TNLQDSTDNDISESESEG 399
```

Search completed: January 29, 2004, 12:39:09
Job time : 14.7105 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	126	100.0	525	4	Q9NQ76	Q9NQ76 homo sapien
2	123	97.6	500	6	Q95KGS	Q95KGS macaca fasc
3	123	97.6	555	6	Q9NQ76	Q9NQ76 macaca fasc
4	123	97.6	555	6	Q9GM13	Q9GM13 macaca fasc
5	81	64.3	433	11	Q92411	Q92411 mus musculus
6	81	64.3	435	11	Q9ES02	Q9ES02 rattus norv
7	81	64.3	441	11	Q8K4L6	Q8K4L6 mus musculus
8	81	64.3	443	11	Q8K3V0	Q8K3V0 rattus norv
9	80	63.5	275	4	Q8NC19	Q8NC19 homo sapien
10	53	42.1	423	4	Q8IU07	Q8IU07 homo sapien
11	52	41.3	320	2	Q52676	Q52676 rhodobacter
12	50	39.7	356	2	Q8L0M4	Q8L0M4 rhodobacter
13	50	39.7	1007	16	Q8EHB9	Q8EHB9 shewanella
14	49.5	39.3	793	5	Q9VVY0	Q9VVY0 drosohilla
15	49	38.9	326	16	Q8NQV1	Q8NQV1 corynebacte
16	49	38.9	339	2	Q8L0L6	Q8L0L6 uncultured

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGDGPFFKD 23
 Db 242 TDLQERGDNDISPFSGDGPFFKD 264

RESULT 2

Q95KG5 ID Q95KG5 PRELIMINARY; PRT; 500 AA.
 AC Q95KG5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 55.6 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Temporal cortex;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB060891; BAB46894.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 500 AA; 55577 MW; 918D265AD8EDC7BC CRC64;

Query Match 97.6%; Score 123; DB 6; Length 500;
 Best Local Similarity 95.7%; Pred. No. 2.4e-10;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGDGPFFKD 23
 Db 218 TDLQERGDNDISPFSGDGPFFKD 240

RESULT 3

Q9N076 ID Q9N076 PRELIMINARY; PRT; 555 AA.
 AC Q9N076;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Matrix extracellular phosphoglycoprotein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB046056; BAB01638.1; -;
 SQ SEQUENCE 555 AA; 61979 MW; CCFEIA98ADAL19E54 CRC64;

Query Match 97.6%; Score 123; DB 6; Length 555;
 Best Local Similarity 95.7%; Pred. No. 2.4e-10;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGDGPFFKD 23
 Db 273 TDLQERGDNDISPFSGDGPFFKD 295

RESULT 4

Q9GM13 ID Q9GM13 PRELIMINARY; PRT; 555 AA.
 AC Q9GM13;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE MEPE protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB050259; BAB17010.1; -;
 SQ SEQUENCE 555 AA; 61950 MW; BD4D47E88038A9E2 CRC64;

Query Match 97.6%; Score 123; DB 6; Length 555;
 Best Local Similarity 95.7%; Pred. No. 2.4e-10;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSGDGPFFKD 23
 Db 273 TDLQERGDNDISPFSGDGPFFKD 295

RESULT 5

Q92411 ID Q92411 PRELIMINARY; PRT; 433 AA.
 AC Q92411;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Matrix extracellular phosphoglycoprotein precursor.
 GN MEPE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21309068; PubMed=11414762;
 RA Argiro L., Desbarats M., Glorieux F.H., Ecarot B.;
 RT "Mepe, the gene encoding a tumor-secreted protein in oncogenic
 RT hypophosphatemic osteomalacia, is expressed in bone";
 RL Genomics 74:342-351 (2001).
 DR EMBL; AF314964; AAK70342.1; -;
 DR MGD; MGI:2137384; Mepe.
 KW Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 433 MATRIX EXTRACELLULAR PHOSPHOGLYCOPROTEIN.
 SQ SEQUENCE 433 AA; 45984 MW; 7CD603CAA8CB41B0 CRC64;

Query Match 64.3%; Score 81; DB 11; Length 433;
 Best Local Similarity 75.0%; Pred. No. 0.00049;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TDLQERGDNDISPFSGDGPFFKD 21
 Db 171 DLQVRGDNDVPPFSGDQHF 190

RESULT 6

Q9ES02 ID Q9ES02 PRELIMINARY; PRT; 435 AA.

```

Q95S02:
AC 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Osteoregulin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20549633; PubMed=10967096;
RA Petersen D.N., Tkalcic G.T., Mansolf A.L., Rivera-Gonzalez R.,
RA Brown T.A.;
RT "Identification of Osteoblast/Osteocyte Factor 45 (OF45), a Bone-
RT specific cDNA Encoding an RGD-containing Protein That Is Highly
RT Expressed in Osteoblasts and Osteocytes.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF260922; AAG33366.1; -.
DR EMBL; AF530558; AAM94403.1; -.
SQ SEQUENCE 435 AA; 46515 MW; D587F82968A26BCB CRC64;

Query Match 64.3%; Score 81; DB 11; Length 435;
Best Local Similarity 75.0%; Pred. No. 0.0005;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPFSGDQPF 21
Db 165 DLLVRGDNVPPFSGDQHF 184

RESULT 7
Q8K4L6 PRELIMINARY; PRT; 441 AA.
AC Q8K4L6;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Osteoregulin.
GN MEPE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Gowen L.C., Petersen D.N., Vail A.L., Stock J., Tkalcic G.T.,
RA Simmons H.A., Childsey-Frink K.L., Ke H., McNeish J., Brown T.A.;
RT "Targeted disruption of the osteoregulin gene results in increased
RT bone density.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF298661; AAM87687.1; -.
DR MGD; MGI:2137384; Mepe.
SQ SEQUENCE 441 AA; 46872 MW; AA1947BFE9F2E300 CRC64;

Query Match 64.3%; Score 81; DB 11; Length 441;
Best Local Similarity 75.0%; Pred. No. 0.0005;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPFSGDQPF 21
Db 179 DLLVRGDNVPPFSGDQHF 198

RESULT 8
Q8K3V0 PRELIMINARY; PRT; 443 AA.
AC Q8K3V0;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Osteoregulin-like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F344;
RA Wang X., Hu B., Wang Y.;
RT "Rattus norvegicus cDNA sequence expressed in B4 cell line (possible
RT subtype of osteoregulin).";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF530559; AAM94404.1; -.
SQ SEQUENCE 443 AA; 47672 MW; 720EEC4C0718FB14 CRC64;

Query Match 64.3%; Score 81; DB 11; Length 443;
Best Local Similarity 75.0%; Pred. No. 0.00051;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPFSGDQPF 21
Db 173 DLLVRGDNVPPFSGDQHF 192

RESULT 9
Q8NC19 PRELIMINARY; PRT; 275 AA.
AC Q8NC19;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90595.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Negahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075076; BAC11386.1; -.
KW Hypothetical protein.
SQ SEQUENCE 275 AA; 31358 MW; AF0B426A671B665C CRC64;

Query Match 63.5%; Score 80; DB 4; Length 275;
Best Local Similarity 93.8%; Pred. No. 0.00042;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISPFSG 16
Db 242 TDLQERGDNDISPFSG 257

RESULT 10
Q8IUP7 PRELIMINARY; PRT; 423 AA.
AC Q8IUP7;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similar to phospholipase B.

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042674; AAI42674.1; -.
SQ SEQUENCE 423 AA; 47354 MW; A20A193491BA780C CRC64;

Query Match 42.1%; Score 53; DB 4; Length 423;
Best Local Similarity 52.2%; Pred. No. 9.4;
Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 TDQERGDNDISPFSGDQPFKD 23
DB 295 TPLNERGDTLTFSEDCFHFS 317

RESULT 11

Q52676 ID Q52676 PRELIMINARY; PRT; 320 AA.

AC Q52676;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE Porin.

GN PORCA.

OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;

OC Rhodobacteraceae; Rhodobacter.

OX NCBI_TaxID=1061;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=37b4;

RX MEDLINE=97149280; PubMed=8996088;

RA Trieschmann M.D., Pattus F., Tadros M.H.;

RT "Molecular characterization and organization of porin from Rhodobacter

RT capsulatus strain 37B4.";

RL Gene 183:61-68(1996).

DR EMBL; U57653; AAB41301.1; -.

DR HSP; P31243; 2POR.

SQ SEQUENCE 320 AA; 33396 MW; B0E1745CB9D86F93 CRC64;

Query Match 41.3%; Score 52; DB 2; Length 320;
Best Local Similarity 58.8%; Pred. No. 9.8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TDQERGDNDISPFSGD 17
DB 112 TDLDRGGNDIPVLTGD 128

RESULT 12

Q8L0M4 ID Q8L0M4 PRELIMINARY; PRT; 356 AA.

AC Q8L0M4;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative formyltetrahydrofolate synthetase (EC 6.3.4.3)

DE (Fragment).

OS uncultured bacterium.

OC Bacteria; environmental samples.

OX NCBI_TaxID=77133;

RN [1]

RP SEQUENCE FROM N.A.

RA Matsui H., Tajima K., Ogata K.;

RT "Diversities of Formyltetrahydrofolate Synthetase Genes Recovered from

RT Rumen Contents.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB085239; BAB96823.1; -.
DR InterPro; IPR000559; Fmtethyd_synth.
DR Pfam; PF01268; FTHFS; 1.
KW Ligase.
FT NON_TER 1 1
FT NON_TER 356 356
SQ SEQUENCE 356 AA; 38922 MW; 485B9F884AC9FD43 CRC64;

Query Match 39.7%; Score 50; DB 2; Length 356;
Best Local Similarity 42.9%; Pred. No. 22;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 DLQERGDNDISPFSGDQPFK 22
DB 80 DLQDRIENILLGITSQK9FR 100

RESULT 13

Q8EHB9 ID Q8EHB9 PRELIMINARY; PRT; 1007 AA.

AC Q8EHB9;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Conserved hypothetical protein.

GN S01309.

OS Shewanella oneidensis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;

OC Alteromonadaceae; Shewanella.

OX NCBI_TaxID=70863;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NR-1.

RX MEDLINE=22297686; PubMed=12368813;

RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,

RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,

RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,

RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,

RA Madupu R., Peterson J.D., Unayam L.A., White O., Wolf A.M.,

RA Vamathevan J., Weidman J., Imbraim M., Lee K., Berry K., Lee C.,

RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,

RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;

RT "Genome sequence of the dissimilatory metal ion-reducing bacterium

RT Shewanella oneidensis.";

RL Nat. Biotechnol. 20:1118-1123(2002).

DR EMBL; AB015574; AAN54374.1; -.

DR TIGR; S01309; -.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 1007 AA; 110092 MW; 8123B725E2CB61F2 CRC64;

Query Match 39.7%; Score 50; DB 16; Length 1007;
Best Local Similarity 52.4%; Pred. No. 71;
Matches 11; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 1 TDLQERGD---NDISPFSGD 17
DB 565 TTKAMRGDAITFNDVPFEGD 585

RESULT 14

Q9VVY0 ID Q9VVY0 PRELIMINARY; PRT; 793 AA.

AC Q9VVY0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE CG14085 protein.

GN CG14085

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

```

RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yeung M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AB003517; AAF49175.1; -.
DR FlyBase; FBgn0036859; CGI4085.
SQ SEQUENCE 793 AA; 88674 MW; FFC0E3DF3E537C41 CRC64;

Query Match 39.3%; Score 49.5; DB 5; Length 793;
Best Local Similarity 61.1%; Pred. No. 65;
Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 7 GNDISPFSG-DGQPFKD 23
Db 693 GNDPSFPFOGQGNKE 710

RESULT 15
ID Q8NOV1 PRELIMINARY; PRT; 326 AA.
AC Q8NOV1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein Cgl1322.
GN CGL1322
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
```

```

RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP005278; BAB967115.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 326 AA; 34230 MW; 9D56CD723B67D686 CRC64;
```

```

Query Match 38.9%; Score 49; DB 16; Length 326;
Best Local Similarity 40.0%; Pred. No. 29;
Matches 10; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

QY 5 ERGDNDISPFSG-----DGQPF 21
Db 163 DSADNPVAPFSSAVDKLEKRDGQPF 187
```

```

Search completed: January 29, 2004, 12:40:18
Job time : 34.8947 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:34:18 ; Search time 28.0263 Seconds
(without alignments)
84.952 Million cell updates/sec

Title: US-09-812-485A-47
Perfect score: 82
Sequence: 1 ERGNDISPFGDQ 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	82	100.0	15 23 AAU93724	Dental product bon
2	82	100.0	15 23 AAQ20377	C-terminal amide
3	82	100.0	23 23 AAU93726	Dental product bon
4	82	100.0	23 23 AAQ20379	C-terminal amide
5	82	100.0	97 23 AAU93681	Dental product bon
6	82	100.0	97 23 AAQ20331	Protein of matrix
7	82	100.0	430 21 AA53812	Amino acid sequenc
8	82	100.0	430 22 AAB62669	Truncated phosphat
9	82	100.0	509 22 AAB82922	Human osteoregulin

10	82	100.0	525 22 AAB82920	Human osteoregulin
11	82	100.0	525 22 AAB62689	Phosphatonin polyp
12	82	100.0	525 23 ABB08526	Human phosphatonin
13	82	100.0	525 23 AA13227	Human oncogenic os
14	82	100.0	540 22 AAB82923	Human osteoregulin
15	82	100.0	556 22 AAB82921	Human osteoregulin
16	77	93.9	40 23 AAU93703	Dental product bon
17	77	93.9	40 23 AAQ20353	Protein of matrix
18	71	86.6	19 23 AA13230	Human OOMI calcium
19	71	86.6	435 22 AAB82918	Rat osteoregulin.
20	71	86.6	441 22 AAB82919	Mouse osteoregulin
21	66	80.5	12 23 AAQ20372	Peptide of a calci
22	60	73.2	15 23 AAU93725	Dental product bon
23	60	73.2	15 23 AAQ20378	C-terminal amide
24	60	73.2	33 23 AAU93710	Dental product bon
25	60	73.2	33 23 AAQ20360	Peptide of matrix
26	60	73.2	275 22 AAM93507	Human polypeptide,
27	55	67.1	15 23 AAU93723	Dental product bon
28	55	67.1	15 23 AAQ20376	C-terminal amide
29	49.5	60.4	40 23 AAU93704	Dental product bon
30	49.5	60.4	40 23 AAQ20354	Protein of matrix
31	49	59.8	30 23 AAU93706	Dental product bon
32	49	59.8	30 23 AAQ20356	Peptide of matrix
33	48.5	59.1	35 23 AAU93705	Dental product bon
34	48.5	59.1	35 23 AAQ20355	Peptide of matrix
35	45	54.9	111 22 AAY72963	Mouse type II inte
36	44.5	54.3	793 22 ABB68973	Drosophila melanog
37	44	53.7	107 22 AAQ12478	Human polypeptide
38	43	52.4	94 22 ABO8384	Novel human diagno
39	43	52.4	648 22 AAU54077	Propionibacterium
40	42	51.2	1228 22 ABB71296	Drosophila melanog
41	41	50.0	109 22 AA10214	Human bone marrow
42	41	50.0	112 22 AAQ04626	Human polypeptide
43	41	50.0	472 23 ABE53556	Human phospholipas
44	41	50.0	530 22 AAQ01956	Human polypeptide
45	41	50.0	1216 23 AAE22860	Human phospholipas

ALIGNMENTS

RESULT 1
AAU93724
ID AAU93724 standard; peptide; 15 AA.

XX AC AAU93724;

XX DT 02-JUL-2002 (first entry)

XX Dental product bone growth enhancing peptide #44.

XX Dental product; toothpaste; mouthwash; dental floss; bone growth;
XX integrin binding motif; RGD; skeletal disease; tooth;
XX alveolar bone growth; osteoblast; odontoblast; osteopathic.

XX OS Synthetic.

XX PN WO200213775-A1.

XX PD 21-FEB-2002.

XX PF 09-AUG-2001; 2001WO-US25101.

XX PR 16-AUG-2000; 2000US-225879P.

XX PA (BIGB-) BIG BEAR BIO INC.

XX PI Yoneda T, Nomizu M, Kumagai Y;

XX DR WPI; 2002-329525/36.

XX Dental product useful for treating skeletal diseases e.g. dental
PT diseases comprises a base material and a compound comprising specific

PT amino acid in a sequence containing the integrin binding motif -
 PS Claim 12; Page 21; 44pp; English.
 CC The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.
 CC The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 82; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ERGNDISPFGDQ 15
 DB 1 ERGNDISPFGDQ 15
 RESULT 2
 AAQ20377
 ID AAO20377 standard; peptide; 15 AA.
 AC AAQ20377;
 DT 31-MAY-2002 (first entry)
 XX
 DE C-terminal amidated synthesised peptide D-00004.
 XX
 KW Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 KW glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 KW alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 KW weakness; D00004.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 15
 FT /note= "C-terminal amide"
 XX
 PN WO200214360-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 14-AUG-2001; 2001WO-US25542.
 XX
 PR 16-AUG-2000; 2000US-0641034.
 PR 19-MAR-2001; 2001US-0812485.
 XX
 PA (BIGB-) BIG BEAR BIO INC.
 XX
 PI Kumagai Y, Blacher RW, Yoneda T;
 XX WPI; 2002-291971/33.
 XX
 PT New peptide compound useful for reducing bone loss, is capable of
 PT enhancing bone growth, and comprises an integrin binding motif,
 PT glycosaminoglycan binding motif or a calcium binding motif -
 XX
 PS Example 1; Page 23; 50pp; English.
 CC
 CC The invention relates to a peptide compound capable of enhancing bone
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino
 CC acids are in D- or L- conformation and the sequence comprises a motif
 CC selected from an integrin binding motif, a glycosaminoglycan binding

CC motif and a calcium binding motif. The peptide of the invention is useful
 CC for reducing bone loss and for reducing renal phosphate excretion in an
 CC individual. The peptide is useful for promoting regeneration of alveolar
 CC bone and/or teeth, and increases the number and activity of odontoblasts
 CC /osteoclasts that help form dental tissues. The peptide is also useful
 CC for treating or preventing a condition associated with skeletal loss or
 CC weakness. This sequence represents a C-terminal amidated synthesised
 CC peptide D-00004 of the invention.
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 82; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ERGNDISPFGDQ 15
 DB 1 ERGNDISPFGDQ 15
 RESULT 3
 AAU93726
 ID AAU93726 standard; peptide; 23 AA.
 AC AAU93726;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Dental product bone growth enhancing peptide #46.
 XX
 KW Dental product; toothpaste; mouthwash; dental floss; bone growth;
 KW integrin binding motif; RGD; skeletal disease; dental disease; tooth;
 KW alveolar bone growth; osteoblast; odontoblast; osteopathic.
 XX
 OS Synthetic.
 XX
 PN WO200213775-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 09-AUG-2001; 2001WO-US25101.
 XX
 PR 16-AUG-2000; 2000US-225879P.
 XX
 PA (BIGB-) BIG BEAR BIO INC.
 XX
 PI Yoneda T, Nomizu M, Kumagai Y;
 XX WPI; 2002-329525/36.
 XX
 PT Dental product useful for treating skeletal diseases e.g. dental
 PT diseases comprises a base material and a compound comprising specific
 PT amino acid in a sequence containing the integrin binding motif -
 XX
 PS Claim 7; Page 21; 44pp; English.
 XX
 CC The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.
 CC The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 82; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQG 15
 DB 5 ERGNDISPFGDQG 19

RESULT 4

AAO20379
 ID AAO20379 standard; peptide; 23 AA.

AC AAO20379;
 DT 31-MAY-2002 (first entry)

XX C-terminal amidated synthesised peptide D-00006.

XX Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 KW glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 KW alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 KW weakness; D00006.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 23 /note= "C-terminal amide"

PN WO200214360-A1.

PD 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US25542.

XX 16-AUG-2000; 2000US-0641034.

XX 19-MAR-2001; 2001US-0812485.

XX (BIGB-) BIG BEAR BIO INC.

PI Kumagai Y, Blacher RW, Yoneda T;

XX WPI; 2002-291971/33.

XX New peptide compound useful for reducing bone loss, is capable of
 PT enhancing bone growth, and comprises an integrin binding motif,
 PT glycosaminoglycan binding motif or a calcium binding motif -
 XX Example 1; Page 15; 50pp; English.

XX The invention relates to a peptide compound capable of enhancing bone
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino
 CC acids are in D- or L- conformation and the sequence comprises a motif
 CC selected from an integrin binding motif, a glycosaminoglycan binding
 CC motif and a calcium binding motif. The peptide of the invention is useful
 CC for reducing bone loss and for reducing renal phosphate excretion in an
 CC individual. The peptide is useful for promoting regeneration of alveolar
 CC bone and/or teeth, and increases the number and activity of odontoblasts
 CC /osteoclasts that help form dental tissues. The peptide is also useful
 CC for treating or preventing a condition associated with skeletal loss or
 CC weakness. This sequence represents a C-terminal amidated synthesised
 CC peptide D-00006 of the invention.

XX Sequence 23 AA;

Query Match 100.0%; Score 82; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQG 15

DB 5 ERGNDISPFGDQG 19

RESULT 5

AAU93681
 ID AAU93681 standard; protein; 97 AA.

XX AAU93681;

XX 02-JUL-2002 (first entry)

XX Dental product bone growth enhancing peptide #1.

XX Dental product; toothpaste; mouthwash; dental floss; bone growth;
 KW integrin binding motif; RGD; skeletal disease; dental disease; tooth;
 KW alveolar bone growth; osteoblast; odontoblast; osteopathic.

OS Synthetic.

XX WO200213775-A1.

XX 21-FEB-2002.

XX 09-AUG-2001; 2001WO-US25101.

XX 16-AUG-2000; 2000US-225879P.

XX (BIGB-) BIG BEAR BIO INC.

PI Yoneda T, Nomizu M, Kumagai Y;

XX WPI; 2002-329525/36.

XX Dental product useful for treating skeletal diseases e.g. dental
 PT diseases comprises a base material and a compound comprising specific
 PT amino acid in a sequence containing the integrin binding motif -
 XX Disclosure; Page 11; 4pp; English.

XX The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.

XX The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.

XX Sequence 97 AA;

Query Match 100.0%; Score 82; DB 23; Length 97;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQG 15

DB 47 ERGNDISPFGDQG 61

RESULT 6

AAO20331
 ID AAO20331 standard; protein; 97 AA.

XX AAO20331;

XX 31-MAY-2002 (first entry)

XX Protein of matrix extracellular phosphoglycoprotein containing RGD #1.

XX Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 KW glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 KW alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 KW weakness; matrix extracellular phosphoglycoprotein.

OS Unidentified.
 XX WO200214360-A1.
 XX
 XX
 XX
 XX 21-FEB-2002.
 XX
 XX 14-AUG-2001; 2001WO-US25542.
 XX
 XX 16-AUG-2000; 2000US-0641034.
 PR 19-MAR-2001; 2001US-0812485.
 XX
 XX (BIGB-) BIG BEAR BIO INC.
 XX
 XX Kumagai Y, Blacher RW, Yoneda T;
 XX WPI; 2002-291971/33.
 DR
 XX
 XX New peptide compound useful for reducing bone loss, is capable of
 PT enhancing bone growth, and comprises an integrin binding motif,
 PT glycosaminoglycan binding motif or a calcium binding motif -
 PT
 XX
 XX Disclosure; Page 11; 50pp; English.
 XX
 CC The invention relates to a peptide compound capable of enhancing bone
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino
 CC acids are in D- or L- conformation and the sequence comprises a motif
 CC selected from an integrin binding motif, a glycosaminoglycan binding
 CC motif and a calcium binding motif. The peptide of the invention is useful
 CC for reducing bone loss and for reducing renal phosphate excretion in an
 CC individual. The peptide is useful for promoting regeneration of alveolar
 CC bone and/or teeth, and increases the number and activity of odontoblasts
 CC /osteoclasts that help form dental tissues. The peptide is also useful
 CC for treating or preventing a condition associated with skeletal loss or
 CC weakness. This sequence represents a protein of a matrix extracellular
 CC phosphoglycoprotein containing an RGD motif of the invention.
 XX
 XX Sequence 97 AA;
 SQ
 Query Match 100.0%; Score 82; DB 23; Length 97;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ERGNDISPFSGDQ 15
 Db |||||
 47 ERGNDISPFSGDQ 61
 RESULT 7
 AAY53812
 ID AAY53812 standard; Protein; 430 AA.
 XX
 AC AAY53812;
 XX
 XX 22-FEB-2000 (first entry)
 DT
 XX
 DE Amino acid sequence of a human phosphatonin polypeptide.
 XX
 KW Human; phosphatonin; Metastatic-tumour Excreted Phosphaturic-Element;
 KW MPE; Na+-dependent phosphate cotransport; vitamin D metabolism;
 KW bone mineralisation; phosphate metabolism related disease;
 KW hyperphosphatemia; renal osteodystrophy; renal dialysis;
 KW secondary hyperparathyroidism; osteitis fibrosa cystica; hypercalcuria;
 KW x-linked hypophosphatemic ricket; hereditary hypophosphatemic ricket;
 KW hypominerallised bone lesion; stunted growth; cystic fibrosis;
 KW oncogenic hypophosphatemics osteomalacia; renal phosphate leakage;
 KW renal osteodystrophy; osteoporosis; vitamin D resistant rickets;
 KW end organ resistance; renal Fanconi syndrome; autosomal rickets;
 KW Paget's disease; kidney failure; renal tubular acidosis; sprue.
 XX
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 8..10
 FT

FT Modified-site
 8..11 /note= "protein kinase C phosphorylation site"
 FT /note= "Casein kinase II phosphorylation site"
 FT 16..21 /note= "myristoylation site"
 FT 40..47 /note= "Tyrosine kinase phosphorylation site"
 FT 77..79 /note= "protein kinase C phosphorylation site"
 FT 118..120 /note= "protein kinase C phosphorylation site"
 FT /note= "protein kinase C phosphorylation site"
 FT 119..124 /note= "myristoylation site"
 FT 139..142 /note= "Casein kinase II phosphorylation site"
 FT 143..148 /note= "myristoylation site"
 FT 152..154 /note= "cell attachment tripeptide"
 FT 161..165 /note= "glycosaminoglycan attachment site"
 FT 177..180 /note= "Casein kinase II phosphorylation site"
 FT 194..197 /note= "Casein kinase II phosphorylation site"
 FT 199..202 /note= "Casein kinase II phosphorylation site"
 FT 203..205 /note= "protein kinase C phosphorylation site"
 FT 224..227 /note= "Casein kinase II phosphorylation site"
 FT 228..231 /note= "Casein kinase II phosphorylation site"
 FT 228..230 /note= "protein kinase C phosphorylation site"
 FT 238..241 /note= "Casein kinase II phosphorylation site"
 FT 266..271 /note= "myristoylation site"
 FT 291..296 /note= "myristoylation site"
 FT 311..313 /note= "protein kinase C phosphorylation site"
 FT 312..314 /note= "protein kinase C phosphorylation site"
 FT 315..320 /note= "myristoylation site"
 FT 319..321 /note= "protein kinase C phosphorylation site"
 FT 325..328 /note= "Casein kinase II phosphorylation site"
 FT 370..373 /note= "amidation site"
 FT 382..386 /note= "Asu-glycosylation site"
 FT 383..387 /note= "Asu-glycosylation site"
 FT 384..386 /note= "protein kinase C phosphorylation site"
 FT 389..394 /note= "myristoylation site"
 FT 403..405 /note= "protein kinase C phosphorylation site"
 FT 405..408 /note= "cAMP and cGMP dependent protien kinase phosphorylation site"
 FT 408..410 /note= "protein kinase C phosphorylation site"
 FT 409..411 /note= "protein kinase C phosphorylation site"
 FT 423..426 /note= "Casein kinase II phosphorylation site"
 FT 425..428

FT Modified-site /note="Casein kinase II phosphorylation site"
 FT 427..430
 FT /note="Casein kinase II phosphorylation site"
 PN W09560017-A2.
 XX 25-NOV-1999.
 PD
 XX
 XX 18-MAY-1999; 99WO-EF03403.
 PF
 XX 18-MAY-1998; 98GB-0010681.
 PR 04-SEP-1998; 98GB-0019387.
 XX
 XX (UNLO) UNIV COLLEGE LONDON.
 PA
 PI Rowe P;
 DR WPI: 2000-053262/04.
 DR N-PSDB; AAZ36447.
 XX
 PT New polypeptides involved in the regulation of phosphate metabolism
 PT useful for diagnosing and treating disorders related to phosphate
 PT metabolism
 PS Claim 6; Fig 8; 136pp; English.
 XX
 CC The present sequence represents a phosphatonin polypeptide (also called
 CC Metastatic-tumour Excreted Phosphaturic-Element (MEPE)). The level of
 CC phosphatonin in a subject modulates Na⁺-dependent phosphate cotransport,
 CC vitamin D metabolism and/or bone mineralisation. The phosphatonin
 CC polypeptides, polynucleotides, vectors and antibodies are used to treat
 CC phosphate metabolism related disease. They are used for treatment of
 CC hyperphosphatemia, or renal osteodystrophy, hyperphosphatemia in renal
 CC dialysis/pre-dialysis, secondary hyperparathyroidism or osteitis fibrosa
 CC cystica, or X-linked hypophosphatemic rickets, hereditary
 CC hypophosphatemic rickets with hypercalcaemia (HHRH), hypominalised
 CC bone lesions, stunted growth in juveniles, oncogenic hypophosphatemic
 CC osteomalacia, renal phosphate leakage, renal osteodystrophy,
 CC osteoporosis, vitamin D resistant rickets, end organ resistance, renal
 CC Fanconi syndrome, autosomal rickets, Paget's disease, kidney failure,
 CC renal tubular acidosis, cystic fibrosis or sprue. The polypeptide
 CC may also be used to manufacture combined preparations for simultaneous,
 CC separate or sequential use for the treatment of phosphate metabolism
 CC disorders. A transformed osteoblast or bone cell line capable of
 CC phosphatonin overexpression is useful for the production of
 CC phosphatonin.
 XX
 SQ Sequence 430 AA;
 Query Match 100.0%; Score 82; DB 21; Length 430;
 Best Local Similarity 100.0%; Pred. No. 2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ERGNDISPFGDQ 15
 DB 151 ERGNDISPFGDQ 165
 RESULT 8
 AAB62669
 ID AAB62669 standard; Protein; 430 AA.
 XX
 AC AAB62669;
 XX
 XX 23-JUL-2001 (first entry)
 DT
 XX Truncated phosphatonin polypeptide (truncated MEPE).
 DE
 XX Metastatic-tumour excreted phosphaturic element; MEPE; phosphatonin;
 KW phosphate; vitamin-D; skeletal formation; mineralization; truncated;
 KW osteopathic; antigit; cytostatic; human.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 161..192
 FT /note="the amino acid residues in this region are
 FT indicated incorrectly in the sequence provided
 FT in the sequence listing"
 XX
 PN W0200132878-A2.
 XX 10-MAY-2001.
 PD
 XX 31-OCT-2000; 2000WO-EP10747.
 PF
 XX 04-NOV-1999; 99US-0434185.
 PR 08-NOV-1999; 99GB-0026424.
 XX
 XX (UNLO) UNIV COLLEGE LONDON.
 PA
 XX Rowe P;
 PI
 DR WPI: 2001-343487/36.
 DR N-PSDB; AAP83759.
 XX
 PT New phosphatonin polypeptide a regulator of phosphate metabolism, for
 PT diagnosing and treating disorders of phosphate, vitamin-D metabolism,
 PT skeletal formation e.g. osteoporosis, Paget's disease, gout
 PS Claim 4; Fig 8; 135pp; English.
 XX
 CC The invention relates to a novel human protein, metastatic-tumour
 CC excreted phosphaturic element (MEPE) or phosphatonin (modulator of
 CC phosphate and vitamin-D metabolism). The phosphatonin polypeptides,
 CC polynucleotides and specific antibodies are useful for treating a
 CC disorder of phosphate or vitamin D metabolism, skeletal formation and
 CC mineralization. Phosphatonins are used to treat hyperphosphatemia, renal
 CC osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica
 CC or gout. It is used to prepare a medicament for treating X-linked
 CC hypophosphatemic rickets, hereditary hypophosphatemic rickets with
 CC hypercalcaemia (HHRH), hypominalised bone lesions, stunted growth in
 CC juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate
 CC leakage, renal osteodystrophy, osteoporosis, vitamin-D resistant rickets,
 CC disease, kidney failure, renal Fanconi syndrome, autosomal rickets, Paget's
 CC disease, kidney failure, renal tubular acidosis, cystic fibrosis or
 CC sprue. Phosphatonin polynucleotides are useful as molecular weight
 CC markers on Southern gels, as diagnostic probes for detecting the presence
 CC of a specific mRNA. Phosphatonin polypeptides are also useful for
 CC identifying agonists and antagonists, compounds which bind to
 CC phosphatonin and drug candidates for therapy of phosphate metabolism
 CC disorders. The present sequence represents a truncated form of
 CC phosphatonin (MEPE).
 XX
 SQ Sequence 430 AA;
 Query Match 100.0%; Score 82; DB 22; Length 430;
 Best Local Similarity 100.0%; Pred. No. 2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ERGNDISPFGDQ 15
 DB 151 ERGNDISPFGDQ 165
 RESULT 9
 AAB82922
 ID AAB82922 standard; Protein; 509 AA.
 XX
 AC AAB82922;
 XX
 XX 21-DEC-2001 (first entry)
 DT
 XX Human osteoregulin (mature polypeptide).
 DE
 XX Osteoregulin; human; bone; homeostasis; adipose; calcification;
 KW

KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
therapy.

XX Homo sapiens.

XX EPI130098-A2.

XX 05-SEP-2001.

XX 27-FEB-2001; 2001EP-0301768.

XX 29-FEB-2000; 2000US-185617P.

XX 22-SEP-2000; 2000US-234500P.

XX (PFIZ) PFIZER PROD INC.

XX Brown TA, De Wet JR, Gowen LC, Hames LM;

XX WPI; 2001-604111/69.

XX N-PSDB; AAH26810.

XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,
adiposity and calcification of atherosclerotic plaques comprises
measuring the activity of osteoregulin -

XX Claim 1; Page 54-55; 90pp; English.

XX The present sequence is that of human osteoregulin mature
polypeptide, i.e. lacking an N-terminal signal sequence.

XX Osteoregulin is a novel protein which plays a role in regulating
bone homeostasis, adiposity, and the calcification of

XX atherosclerotic plaques. 2 Splice variants of human osteoregulin
were identified (see also AAH82923). The invention provides novel

XX osteoregulin proteins, nucleic acids which encode them, vectors,
antibodies, host cells which express heterologous osteoregulins, and

XX animal cells and mammals with a targeted disruption of an
osteoregulin gene. The invention also provides screening assays

XX to identify modulators of osteoregulin activity as well as methods
of treating mammals for diseases or disorders associated with

XX osteoregulin activity. The modulators of activity may be useful
in the manufacture of a medicament for, as well as for treating, a

XX mammal in need of regulation of bone mass and/or density, adiposity,
vascular flexibility, and/or atherosclerotic plaque calcification

XX (claimed), for treating and preventing osteoporosis, and for
stimulating bone repair and regeneration.

XX Sequence 509 AA;

SQ Query Match 100.0%; Score 82; DB 22; Length 509;

ID Best Local Similarity 100.0%; Pred. No. 2.4e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPSPSGDQ 15

DB 230 ERGNDISPSPSGDQ 244

RESULT 10

AAH82920

XX ID AAB82920 standard; Protein; 525 AA.

XX AC AAB82920;

XX DT 21-DEC-2001 (first entry)

XX DE Human osteoregulin.

XX Osteoregulin; human; bone; homeostasis; adipose; calcification;

XX atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
therapy.

XX Homo sapiens.

XX OS

6

PH Key Location/Qualifiers

FT Peptide 1..37

FT Protein /label= Signal_peptide

FT /label= Mature_protein

PN EPI130098-A2.

XX 05-SEP-2001.

XX 27-FEB-2001; 2001EP-0301768.

XX 29-FEB-2000; 2000US-185617P.

XX 22-SEP-2000; 2000US-234500P.

XX (PFIZ) PFIZER PROD INC.

XX Brown TA, De Wet JR, Gowen LC, Hames LM;

XX WPI; 2001-604111/69.

XX N-PSDB; AAH26808.

XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,
adiposity and calcification of atherosclerotic plaques comprises
measuring the activity of osteoregulin -

XX Claim 1; Page 45-47; 90pp; English.

XX The present sequence is that of human osteoregulin, a novel protein
which plays a role in regulating bone homeostasis, adiposity, and
the calcification of atherosclerotic plaques. The sequence is

XX predicted from the nucleotide sequence (see AAH26808) of isolated
cDNA. A splice variant of human osteoregulin was also

XX identified (see AAB82921). The invention provides novel
osteoregulin proteins, nucleic acids which encode them, vectors,

XX antibodies, host cells which express heterologous osteoregulins, and
animal cells and mammals with a targeted disruption of an

XX osteoregulin gene. The invention also provides screening assays
to identify modulators of osteoregulin activity as well as methods

XX of treating mammals for diseases or disorders associated with
osteoregulin activity. The modulators of activity may be useful

XX in the manufacture of a medicament for, as well as for treating, a
mammal in need of regulation of bone mass and/or density, adiposity,

XX vascular flexibility, and/or atherosclerotic plaque calcification
(claimed), for treating and preventing osteoporosis, and for

XX stimulating bone repair and regeneration.

SQ Sequence 525 AA;

Query Match 100.0%; Score 82; DB 22; Length 525;

ID Best Local Similarity 100.0%; Pred. No. 2.5e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPSPSGDQ 15

DB 246 ERGNDISPSPSGDQ 260

RESULT 11

AAH2689

XX ID AAB62689 standard; Protein; 525 AA.

XX AC AAB62689;

XX DT 23-JUL-2001 (first entry)

XX DE Phosphatonin polypeptide (MEPE).

XX Metastatic-tumour excreted phosphaturic element; MEPE; phosphatonin;
phosphate; vitamin-D; skeletal formation; mineralization; truncated;
osteopathic; antitumor; cytostatic; human.

XX Homo sapiens.

XX OS

```

XX PN WO200132878-A2.
XX PD 10-MAY-2001.
XX PF 31-OCT-2000; 2000WO-EP10747.
XX PR 04-NOV-1999; 99US-0434185.
XX PR 08-NOV-1999; 99GB-0026424.
XX PA (UNLO ) UNIV COLLEGE LONDON.
XX PI Rowe P;
XX DR WPI; 2001-343487/36.
XX DR N-PSDB; AAF83764.
XX PT New phosphatonin polypeptide a regulator of phosphate metabolism, for
XX PT diagnosing and treating disorders of phosphate, vitamin-D metabolism,
XX PT skeletal formation e.g. osteoporosis, Paget's disease, gout
XX PS Claim 4; Page 133-134; 135pp; English.
XX CC The invention relates to a novel human protein, metastatic-tumour
XX CC excreted phosphaturic element (MEPE) or phosphatonin (modulator of
XX CC phosphate and vitamin-D metabolism). The phosphatonin polypeptides,
XX CC polynucleotides and specific antibodies are useful for treating a
XX CC disorder of phosphate or vitamin D metabolism, skeletal formation and
XX CC mineralization. Phosphatonins are used to treat hyperphosphatemia, renal
XX CC osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica
XX CC or gout. It is used to prepare a medicament for treating X-linked
XX CC hypophosphatemic rickets, hereditary hypophosphatemic rickets with
XX CC hypercalcaemia (HHRH), hypominealized bone lesions, stunted growth in
XX CC juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate
XX CC leakage, renal osteodystrophy, osteoporosis, vitamin-D resistant rickets,
XX CC end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's
XX CC disease, kidney failure, renal tubular acidosis, cystic fibrosis or
XX CC sprue. Phosphatonin polynucleotides are useful as molecular weight
XX CC markers on Southern gels, as diagnostic probes for detecting the presence
XX CC of a specific mRNA. Phosphatonin polypeptides are also useful for
XX CC identifying agonists and antagonists, compounds which bind to
XX CC phosphatonin and drug candidates for therapy of phosphate metabolism
XX CC disorders. The present sequence represents the amino acid sequence of
XX CC the entire phosphatonin (MEPE).
XX SQ Sequence 525 AA;
Query Match 100.0%; Score 82; DB 22; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGDNDISPFSGDQG 15
DB 246 ERGDNDISPFSGDQG 260

RESULT 12
AB080526
ID ABB080526 standard; protein; 525 AA.
XX AC ABB080526;
XX DT 23-MAY-2002 (first entry)
XX FT Human phosphatonin.
XX DE Human phosphatonin.
XX KW Human phosphatonin; cytostatic; antidiabetic; antiinflammatory;
XX KW hyperphosphatemia; arteriosclerosis; heart failure;
XX KW diabetic renal disease; kidney failure; cystic fibrosis.
XX OS Homo sapiens.
XX PN WO200198495-A1.

XX PD 27-DEC-2001.
XX PF 20-JUN-2001; 2001WO-JP05263.
XX PR 21-JUN-2000; 2000JP-0191088.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Kurokawa T, Yamada T, Morimoto S;
XX DR WPI; 2002-139791/18.
XX DR N-PSDB; ABA99159, ABA99160.
XX PT Phosphatonin of human origin and DNA encoding it for diagnosis and
XX PT treatment of diseases associated with disorders of phosphate
XX PT metabolism, e.g., hyperphosphatemia, arteriosclerosis, heart failure,
XX PT diabetic renal disease and kidney failure -
XX PS Claim 1; Fig 1-4; 130pp; Japanese.
XX CC This invention relates to human phosphatonin which has
XX CC the activity of lowering blood phosphate and increasing urinary
XX CC phosphate. The proteins are cytostatic, antidiabetic and
XX CC antiinflammatory in their action. Phosphatonin down-regulates
XX CC sodium-dependent phosphate transport in the kidney, it down-regulates
XX CC 25-hydroxy-vitamin D3-1alpha-hydroxylase in the kidney and up-regulates
XX CC 25-hydroxy-vitamin D3-24-hydroxylase in the kidney. Phosphatonin is
XX CC useful in the diagnosis, treatment and prevention of phosphate
XX CC metabolism related diseases such as hyperphosphatemia, arteriosclerosis,
XX CC heart failure, diabetic renal disease, kidney failure, acute coronary
XX CC disease and cystic fibrosis. This sequence represents human
XX CC phosphatonin.
XX SQ Sequence 525 AA;
Query Match 100.0%; Score 82; DB 23; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGDNDISPFSGDQG 15
DB 246 ERGDNDISPFSGDQG 260

RESULT 13
AAE13227
ID AAE13227 standard; Protein; 525 AA.
XX AC AAE13227;
XX DT 12-FEB-2002 (first entry)
XX DE Human oncogenic osteomalacia-related protein-1. (OOM-1).
XX KW Human; oncogenic osteomalacia-related protein 1; OOM1; gene therapy;
XX KW phosphate homeostasis-related disease; rickets; osteomalacia; cardiac;
XX KW rhabdomyolysis; cardiomyopathy; tumoral calcinosis; renal failure;
XX KW bone mineralisation; phosphaturia; cellular process.
XX OS Homo sapiens.
XX FT Key
XX FT Peptide
XX FT Protein
XX FT Domain
XX FT Domain
XX FT Region
XX FT Location/Qualifiers
XX FT 1..16
XX FT /label= Signal_peptide
XX FT /note= "Mature oncogenic osteomalacia-related protein-1
XX FT (OOM-1)"
XX FT 130..142
XX FT /note= "Calcium binding motif"
XX FT 235..258
XX FT /note= "Calcium binding motif"
XX FT 247..249

```

FT /note= "Cell attachment site"
FT 264..275
FT /note= "Calcium binding motif"
FT 412..424
FT /note= "Calcium binding motif"
FT 442..455
FT /note= "Potential PKA phosphorylation site"
FT 477..481
FT /note= "Glycosylation site"
FT 500..503
FT /note= "Bacterial extracellular solute-binding
FT protein motif"
XX WO200172826-A2.
XX
XX
XX 04-OCT-2001.
XX
XX 22-MAR-2001; 2001WO-US09289.
XX
XX 24-MAR-2000; 2000US-191786P.
XX 19-OCT-2000; 2000US-241598P.
XX
XX (GENZ) GENZYME CORP.
XX (UYJO) UNIV JOHNS HOPKINS.
XX
XX Schiavi S, Madden S, Manavalan P, Levine MDM, Jan De Beur S;
XX WPI; 2002-010720/01.
XX
XX New polynucleotide encoding a polypeptide useful for identifying
FT polynucleotide expression or polypeptide activity modulators used for
FT treating oncogenic osteomalacia, comprises the oncogenic
FT osteomalacia-related gene -
XX
XX
PS Claim 2; Fig 2; 65pp; English.
XX
XX The invention relates to oncogenic osteomalacia-related protein 1
CC designated OOM1 (oncogenic osteomalacia-related factor) and its DNA
CC molecule. OOM1 protein is useful for treating a phosphate homeostasis-
CC related disease such as X-linked hypophosphataemia rickets, oncogenic
CC osteomalacia, rhabdomyolysis, cardiomyopathy, tumoural calcinosis and
CC renal failure. OOM1 proteins and DNA's are useful for modulating the
CC phenotype of a neoplastic cell associated with oncogenic osteomalacia;
CC modulating bone mineralisation; treating a disease characterised by
CC abnormal bone mineralisation; and treating an oncogenic osteomalacia-
CC associated symptom such as hypophosphataemia, phosphaturia, low serum
CC concentrations of 1,25-dihydroxy vitamin D or osteomalacia. OOM1 is
CC useful for modulating renal phosphate transport, which involves altering
CC the activity of OOM1 within a cell and then administering serine protease
CC 11 to the cell. OOM1 DNA is useful for generating non-human transgenic
CC animal models, and for searching and identifying single nucleotide
CC polymorphisms which are mutants, variants of the gene in human
CC population. OOM1 is useful as an immunogen to produce antibodies against
CC OOM1 and in vitro assays to screen for agents or drugs, which modulate
CC cellular processes. The present sequence is human oncogenic
CC osteomalacia-related protein-1 (OOM-1).
CC Note: The present sequence states that this sequence is encoded by the
CC DNA sequence shown in SEQ ID NO: 1 (AAB21889) of the specification.
CC However this does not appear to be the case.
XX
SQ Sequence 525 AA;

Query Match 100.0%; Score 82; DB 23; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQG 15
Db 246 ERGNDISPFGDQG 260
|||||

RESULT 14
AAB82923

ID AAB82923 standard; Protein; 540 AA.
XX
AC AAB82923;
XX
DT 21-DEC-2001 (first entry)
XX
DE Human osteoregulin (mature polypeptide).
XX
DE Osteoregulin; human; bone; homeostasis; adipose; calcification;
KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
KW therapy.
XX
OS Homo sapiens.
XX
PN EP1130098-A2.
XX
PD 05-SEP-2001.
XX
PF 27-FEB-2001; 2001EP-0301768.
XX
PR 29-FEB-2000; 2000US-185617P.
XX
PR 22-SEP-2000; 2000US-234500P.
XX
PA (PFIZ) PFIZER PROD INC.
XX
PI Brown TA, De Wet JR, Gowen LC, Hames LM;
XX WPI; 2001-604111/69.
DR N-PSDB; AAB26811.
XX
XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,
PT adiposity and calcification of atherosclerotic plaques comprises
PT measuring the activity of osteoregulin -
XX
PS Claim 1; Page 59-61; 90pp; English.
XX
XX The present sequence is that of human osteoregulin mature
CC polypeptide, i.e. lacking an N-terminal signal sequence.
CC Osteoregulin is a novel protein which plays a role in regulating
CC bone homeostasis, adiposity, and the calcification of
CC atherosclerotic plaques. 2 Splice variants of human osteoregulin
CC were identified (see also AAB82922). The invention provides novel
CC osteoregulin proteins, nucleic acids which encode them, vectors,
CC antibodies, host cells which express heterologous osteoregulins, and
CC animal cells and mammals with a targeted disruption of an
CC osteoregulin gene. The invention also provides screening assays
CC to identify modulators of osteoregulin activity as well as methods
CC of treating mammals for diseases or disorders associated with
CC osteoregulin activity. The modulators of activity may be useful
CC in the manufacture of a medicament for, as well as for treating, a
CC mammal in need of regulation of bone mass and/or density, adiposity,
CC vascular flexibility, and/or atherosclerotic plaque calcification
CC (claimed), for treating and preventing osteoporosis, and for
CC stimulating bone repair and regeneration.
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 82; DB 22; Length 540;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERGNDISPFGDQG 15
Db 261 ERGNDISPFGDQG 275
|||||

RESULT 15
AAB82921
ID AAB82921 standard; Protein; 556 AA.
XX
AC AAB82921;
XX
DT 21-DEC-2001 (first entry)


```
XX Human osteoregulin.
DE
XX Osteoregulin; human; bone; homeostasis; adipose; calcification;
KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
KW therapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..37
FT /label= Signal_peptide
FT Protein 38..556
FT /label= Mature_protein
FT Misc-difference 57
FT /note= "encoded by TCA"
XX
XX EP1130098-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 27-FEB-2001; 2001EP-0301768.
PF
XX
XX 29-FEB-2000; 2000US-185617P.
PR
XX 22-SEP-2000; 2000US-234500P.
PR
XX
XX (PFIZ ) PFIZER PROD INC.
PA
XX
XX Brown TA, De Wet JR, Gowen LC, Hames LM;
PI
XX WPI; 2001-604111/69.
XX N-PSDB; AAH26809.
DR
XX
XX Novel osteoregulin polypeptide useful for regulating bone homeostasis,
PT adiposity and calcification of atherosclerotic plaques comprises
PT measuring the activity of osteoregulin -
XX
XX Claim 1; Page 48-49; 90pp; English.
XX
XX The present sequence is that of human osteoregulin, a novel protein
CC which plays a role in regulating bone homeostasis, adiposity, and
CC the calcification of atherosclerotic plaques. The sequence is
CC predicted from the nucleotide sequence (see AAH26809) of isolated
CC osteoblast cDNA. A splice variant of human osteoregulin was also
CC identified (see AAB82920). The invention provides novel
CC osteoregulin proteins, nucleic acids which encode them, vectors,
CC antibodies, host cells which express heterologous osteoregulins, and
CC animal cells and mammals with a targeted disruption of an
CC osteoregulin gene. The invention also provides screening assays
CC to identify modulators of osteoregulin activity as well as methods
CC of treating mammals for diseases or disorders associated with
CC osteoregulin activity. The modulators of activity may be useful
CC in the manufacture of a medicament for, as well as for treating, a
CC mammal in need of regulation of bone mass and/or density, adiposity,
CC vascular flexibility, and/or atherosclerotic plaque calcification
CC (claimed), for treating and preventing osteoporosis, and for
CC stimulating bone repair and regeneration.
XX
XX Sequence 556 AA;
SQ
Query Match 100.0%; Score 82; DB 22; Length 556;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ERGNDISPESGDGQ 15
Db 277 ERGNDISPESGDGQ 291
Search completed: January 29, 2004, 12:38:34
Job time : 28.0263 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 12:37:19 ; Search time 11.4474 Seconds
(without alignments)
55.442 Million cell updates/sec

Title: US-09-812-485A-47
Perfect score: 82
Sequence: 1 ERGNDISFFSGDQ 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/aaa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/aaa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/aaa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/aaa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/aaa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	50.0	2039	US-09-077-098A-7	Sequence 7, Appli
2	41	50.0	2042	US-09-077-098A-6	Sequence 6, Appli
3	40	48.8	302	US-09-252-991A-21875	Sequence 21875, A
4	40	48.8	795	US-09-370-807-6	Sequence 6, Appli
5	40	48.8	795	US-09-921-259-6	Sequence 6, Appli
6	39	47.6	101	US-09-252-991A-22899	Sequence 22899, A
7	39	47.6	425	US-09-252-991A-31834	Sequence 31834, A
8	38.5	47.0	469	US-08-704-711A-16	Sequence 16, Appli
9	38.5	47.0	469	US-08-448-489-12	Sequence 12, Appli
10	38.5	47.0	469	US-09-521-220-16	Sequence 16, Appli
11	38.5	47.0	469	US-09-391-104-23	Sequence 23, Appli
12	38	46.3	169	US-09-252-991A-26193	Sequence 26193, A
13	38	46.3	186	US-09-328-352-7703	Sequence 7703, Ap
14	38	46.3	3038	US-08-450-332-2	Sequence 2, Appli
15	38	46.3	3038	US-08-637-640-2	Sequence 2, Appli
16	38	46.3	3038	US-09-004-406C-2	Sequence 2, Appli
17	37	45.1	232	US-09-134-001C-5220	Sequence 5220, Ap
18	37	45.1	245	US-09-252-991A-30321	Sequence 30321, A
19	37	45.1	256	PCT-US96-01314-55	Sequence 55, Appli
20	37	45.1	301	US-08-355-844-2	Sequence 2, Appli
21	37	45.1	301	PCT-US95-16126-2	Sequence 2, Appli
22	37	45.1	313	US-09-252-991A-23810	Sequence 23810, A
23	37	45.1	317	US-09-252-991A-31126	Sequence 31126, A
24	37	45.1	394	US-07-637-870-1	Sequence 1, Appli
25	37	45.1	394	US-07-637-399-1	Sequence 1, Appli
26	37	45.1	394	US-07-640-476-5	Sequence 5, Appli
27	37	45.1	394	US-08-112-703-1	Sequence 1, Appli

28	37	45.1	394	6	5290690-2	Patent No. 5290690
29	37	45.1	683	4	US-09-252-991A-19718	Sequence 19718, A
30	37	45.1	799	1	US-08-054-077C-2	Sequence 2, Appli
31	37	45.1	829	4	US-09-252-991A-27150	Sequence 27150, A
32	37	45.1	1198	4	US-09-199-637A-405	Sequence 405, App
33	36.5	44.5	1257	1	US-08-340-428B-49	Sequence 49, Appl
34	36	43.9	40	3	US-09-004-406C-16	Sequence 16, Appl
35	36	43.9	106	3	US-09-113-977C-42	Sequence 42, Appl
36	36	43.9	106	3	US-09-113-977C-43	Sequence 42, Appl
37	36	43.9	106	4	US-09-351-048A-42	Sequence 42, Appl
38	36	43.9	106	4	US-09-351-048A-43	Sequence 43, Appl
39	36	43.9	107	3	US-09-113-977C-41	Sequence 41, Appl
40	36	43.9	107	4	US-09-351-048A-41	Sequence 41, Appl
41	36	43.9	257	1	US-08-451-947-10	Sequence 10, Appl
42	36	43.9	257	2	US-08-424-826A-10	Sequence 10, Appl
43	36	43.9	257	3	US-08-928-694-10	Sequence 10, Appl
44	36	43.9	257	4	US-08-450-842-10	Sequence 10, Appl
45	36	43.9	257	4	US-08-451-390-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-077-098A-7
; Sequence 7, Application US/09077098A
; Patent No. 6544519
; GENERAL INFORMATION:
; APPLICANT: TOKUNAGA, Eiji
; SAKAGUCHI, Masaaki
; MATSUO, Kazuo
; HAMADA, Fukusaburo
; TOKIYOSHI, Sachio
; TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
; PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,098A
FILING DATE: 19-May-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP97/03222
FILING DATE: 12-SEP-1997
APPLICATION NUMBER: JP 27,148/1996
FILING DATE: 19-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: KORNBAU, Anne M.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TOKUNAGA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2039 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-077-098A-7

Query Match 50.0%; Score 41; DB 4; Length 2039;
 Best Local Similarity 80.0%; Pred. No. 4.1e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 DNDISPFSGD 13
 DB 1599 DGDISPTSGD 1608

RESULT 2
 US-09-077-098A-6
 ; Sequence 6, Application US/09077098A
 ; Patent No. 6544519
 ; GENERAL INFORMATION:
 ; APPLICANT: TOKUNAGA, Ei-ji
 ; SAKAGUCHI, Masashi
 ; MATSUO, Kazuo
 ; HAMADA, Fukusaburo
 ; TOKIYOSHI, Sachio
 ; TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
 ; PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWDY AND NEIMARK
 ; STREET: 624 Ninth Street, N.W., Suite 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20001
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/077,098A
 ; FILING DATE: 19-May-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP97/03222
 ; FILING DATE: 12-SEP-1997
 ; APPLICATION NUMBER: JP 27,148/1996
 ; FILING DATE: 19-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KORNBAU, Anne M.
 ; REGISTRATION NUMBER: 25,618
 ; REFERENCE/DOCKET NUMBER: TOKUNAGA-1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; INFORMATION FOR SEQ ID NO: 6:
 ; LENGTH: 2042 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 US-09-077-098A-6

Query Match 50.0%; Score 41; DB 4; Length 2042;
 Best Local Similarity 80.0%; Pred. No. 4.1e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 DNDISPFSGD 13
 DB 1591 DGDISPTSGD 1600

RESULT 3
 US-09-252-991A-21875
 ; Sequence 21875, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 21875
 ; LENGTH: 302
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-21875

Query Match 48.8%; Score 40; DB 4; Length 302;
 Best Local Similarity 46.7%; Pred. No. 76;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 ERGNDISPFSGDQ 15
 DB 137 QRGDIDVAILFGDGR 151

RESULT 4
 US-09-370-807-6
 ; Sequence 6, Application US/09370807
 ; Patent No. 6297034
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Falco, S. Carl
 ; APPLICANT: Rafalski, J. Antoni
 ; APPLICANT: Sakai, Hajime
 ; TITLE OF INVENTION: N-End Rule Pathway Enzymes
 ; FILE REFERENCE: BB-1199
 ; CURRENT APPLICATION NUMBER: US/09/370,807
 ; CURRENT FILING DATE: 1999-08-09
 ; EARLIER APPLICATION NUMBER: 60/096,225
 ; EARLIER FILING DATE: August 12, 1998
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 6
 ; LENGTH: 795
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-370-807-6

Query Match 48.8%; Score 40; DB 3; Length 795;
 Best Local Similarity 60.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GNDISPFSG 12
 DB 427 GDRELNPFG 436

RESULT 5
 US-09-921-259-6
 ; Sequence 6, Application US/09921259
 ; Patent No. 6465234
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Falco, S. Carl
 ; APPLICANT: Rafalski, J. Antoni
 ; APPLICANT: Sakai, Hajime
 ; TITLE OF INVENTION: N-End Rule Pathway Enzymes
 ; FILE REFERENCE: BB-1199
 ; CURRENT APPLICATION NUMBER: US/09/921,259
 ; CURRENT FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: 60/096,225
 ; PRIOR FILING DATE: August 12, 1998

1 NUMBER OF SEQ ID NOS: 16
2 SOFTWARE: Microsoft Office 97
3 SEQ ID NO 6
4 LENGTH: 795
5 TYPE: PRT
6 ORGANISM: Zea mays
7 US-09-921-259-6

Query Match 48.8%; Score 40; DB 4; Length 795;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GNDISPFSG 12
DB 427 GDRELNPFSG 436

RESULT 6
US-09-252-991A-22899
; Sequence 22899, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22899
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22899

Query Match 47.6%; Score 39; DB 4; Length 101;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 DNDISPFSG 13
DB 17 DGLKPFSG 26

RESULT 7
US-09-252-991A-31834
; Sequence 31834, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31834
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31834

Query Match 47.6%; Score 39; DB 4; Length 425;
Best Local Similarity 53.3%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ERGDNDISPFSGDG 15
DB 65 EQGRNDASESTGGG 79

RESULT 8
US-08-704-711A-16
; Sequence 16, Application US/08704711A
; Patent No. 6114159
; GENERAL INFORMATION:
; APPLICANT: WILL. Horet
; APPLICANT: HINZMANN, Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,711A
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DE95/00357
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-704-711A-16

Query Match 47.0%; Score 38.5; DB 3; Length 469;
Best Local Similarity 64.3%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 2 RGDNDISPFSGDG 14
DB 165 RGDHRDNDSPFDGFG 178

RESULT 9
US-08-448-489-12
; Sequence 12, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi

```

; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Known Member of
; OTHER INFORMATION: Matrix Metalloproteinase Family
US-08-448-489-12

Query Match          47.0%; Score 38.5; DB 3; Length 469;
Best Local Similarity 64.3%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      2 RGDN-DISPFSGDG 14
Db      165 RGDHRDNSPFDGPG 178

RESULT 10
US-09-521-220-16
; Sequence 16, Application US/09521220
; Patent No. 6399348
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; HINZMANN, Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/521,220
; FILING DATE: 08-Mar-2000
; CLASSIFICATION: <Unknown>
; 21-OCT-1994
; 17-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,711
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-521-220-16

Query Match          47.0%; Score 38.5; DB 4; Length 469;
Best Local Similarity 64.3%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      2 RGDN-DISPFSGDG 14
Db      165 RGDHRDNSPFDGPG 178

RESULT 11
US-09-391-104-23
; Sequence 23, Application US/093911104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-391-104-23

Query Match          47.0%; Score 38.5; DB 4; Length 469;
Best Local Similarity 64.3%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      2 RGDN-DISPFSGDG 14
Db      165 RGDHRDNSPFDGPG 178

RESULT 12
US-09-252-991A-26193
; Sequence 26193, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26193
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26193

Query Match          46.3%; Score 38; DB 4; Length 169;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

Qy 3 GDNDISPFGDG 14
 Db 6 GGSDISPDSGAG 17

RESULT 13

US-09-328-352-7703
 ; Sequence 7703, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 7703
 ; LENGTH: 186
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-7703

Query Match 46.3%; Score 38; DB 4; Length 186;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GDNDISPFS 11
 Db 77 GETDVSFSS 85

RESULT 14

US-08-450-332-2
 ; Sequence 2, Application US/08450332
 ; Patent No. 5744350
 ; GENERAL INFORMATION:
 ; APPLICANT: VINCI, VICTOR A.
 ; APPLICANT: CONDER, MICHAEL J.
 ; APPLICANT: MCADA, PHYLLIS C.
 ; APPLICANT: REEVES, CHRISTOPHER D.
 ; APPLICANT: DAVIS, CHARLES R.
 ; APPLICANT: HENDRICKSON, LEE E.
 ; APPLICANT: RAMOSEK, JOHN
 ; TITLE OF INVENTION: DNA ENCODING TRIOL POLYKETIDE SYNTHASE
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CHRISTINE E. CARTY
 ; STREET: 126 E. LINCOLN AVENUE, P.O. Box 2000
 ; CITY: RAHWAY
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0907
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/450,332
 ; FILING DATE: 25-MAY-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CARTY, CHRISTINE E.
 ; REGISTRATION NUMBER: 36,099
 ; REFERENCE/DOCKET NUMBER: 19076CA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 908-594-6734
 ; TELEFAX: 908-594-4720
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3038 amino acids
 ; TYPE: amino acid

Query Match 46.3%; Score 38; DB 4; Length 186;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: TPKS Protein
 US-08-450-332-2

Query Match 46.3%; Score 38; DB 1; Length 3038;
 Best Local Similarity 63.6%; Pred. No. 1.9e+03;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ERGNDISPFS 11
 Db 878 ERGKNDLSFS 888

RESULT 15

US-08-637-640-2
 ; Sequence 2, Application US/08637640
 ; Patent No. 5849541
 ; GENERAL INFORMATION:
 ; APPLICANT: VINCI, VICTOR A.
 ; APPLICANT: CONDER, MICHAEL J.
 ; APPLICANT: MCADA, PHYLLIS C.
 ; APPLICANT: REEVES, CHRISTOPHER D.
 ; APPLICANT: DAVIS, CHARLES R.
 ; APPLICANT: HENDRICKSON, LEE E.
 ; APPLICANT: RAMOSEK, JOHN
 ; TITLE OF INVENTION: DNA ENCODING TRIOL POLYKETIDE SYNTHASE
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CHRISTINE E. CARTY
 ; STREET: P.O. BOX 2000; 126 E. LINCOLN AVENUE
 ; CITY: RAHWAY
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/637,640
 ; FILING DATE: 23-AUG-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/148,132
 ; FILING DATE: 01-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CARTY, CHRISTINE E.
 ; REGISTRATION NUMBER: 36,099
 ; REFERENCE/DOCKET NUMBER: 19076
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 908-594-6734
 ; TELEFAX: 908-594-4720
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3038 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: TPKS Protein
 ; US-08-637-640-2

Query Match 46.3%; Score 38; DB 2; Length 3038;
 Best Local Similarity 63.6%; Pred. No. 1.9e+03;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ERGNDISFFS 11
|||
Db 878 ERGNDLDSFS 888

Search completed: January 29, 2004, 12:41:37
Job time : 12.4474 secs